

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 11:44:56 ; Search time 35 Seconds
(without alignments)
22.843 Million cell updates/sec

Title: US-09-919-703-1
Perfect score: 30
Sequence: 1 SVDVEY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	30	100.0	6 20	AAV25009 Streptokinase deri
2	30	100.0	6 23	ABB80001 Streptokinase deri
3	30	100.0	16 23	ABB80005 Streptokinase deri
4	30	100.0	18 20	AAV25013 Streptokinase deri
5	30	100.0	20 20	AAV25014 Streptokinase deri
6	30	100.0	20 23	ABB80006 Streptokinase deri
7	30	100.0	21 20	AAV25015 Streptokinase deri
8	30	100.0	21 23	ABB80007 Streptokinase deri
9	30	100.0	71 20	AAV01550 Antigenic epitope
10	30	100.0	146 20	AAV01558 Peptide correspond

ALIGNMENTS

RESULT 1

AAV25009	AAV25009	standard; peptide; 6 AA.
ID	AAV25009	standard; peptide; 6 AA.
XX	XX	
AC	AAV25009;	
XX	XX	
DT	23-AUG-1999	(first entry)
XX	XX	
DE	DE	Streptokinase derived peptide 1 for amelioration of cell death.
XX	XX	
XX	XX	Streptokinase; cell death; amelioration; treatment; disease; aging; AIDS; cellular differentiation; physical insult; trauma; anoxia; hyperthermia; hypothermia; brain; spinal cord; kidney; heart; lung; liver; skin; HIV; viral disorder; hepatitis; retroviral; infection; encephalitis; FALS; neurodegenerative disorder; Parkinson's disease; Alzheimer's disease; Huntington's disease; cerebellar degeneration; cardiovascular disease; familial amyotrophic lateral sclerosis; atherosclerosis; heart failure; infarction; heart disease; cardiomyopathy; hypertensive; myocardial; reperfusion injury; immune disease; rheumatoid arthritis; renal; retinal; systemic lupus erythematosus; insulin-dependent; diabetes mellitus; pernicious anaemia; dermatomyositis; erythema nodosum; myasthenia gravis; Sjogren's syndrome; temporal arthritis; autoimmune polyarthritides; Wegener's granulomatosis; glomerulonephritis; anti-phospholipid syndrome; neoplastic disorder; leukemia; sarcoma; myeloma; carcinoma; neuroma; melanoma; cancer; breast; colon; cervix; prostate; Hodgkin's disease; non-Hodgkin's lymphoma; inflammatory disorder; stroke; ischemia; spinal cord; toxic insult; pulmonary; macular degeneration; cataract; pancreatitis; Crohn's disease; ulcerative colitis; motor neuron disease; Guillan Bare Syndrome; demyelinating disease; bypass surgery; chemotherapy; clozapine; AZT; anthracycline.
XX	XX	Synthetic.
OS	OS	

Recombinant strept
Peptide correspond
Streptokinase from
Streptokinase (1-3
Recombinant strept
Truncated Met stre
Streptokinase (1-3
Streptokinase (1-3
Streptokinase (1-3
Amino acid sequenc
Streptokinase(1-37
Amino acid sequenc
Streptokinase (1-3
Amino acid sequenc
Streptokinase frag
Streptokinase. St
S.equisimilis stre
Native streptokina
Streptococcus equi
Streptococcus equi
Streptococcus (SK)
De-immunised strep
S. equisimilis str
Wild type streptok
Altered streptokin
Mutant streptokina
Streptococcus plas
Streptococcus sp.
Sequence encoded b
Streptokinase A fr
Streptokinase G pr

PN US5917013-A.
PD 29-JUN-1999.
XX 05-DEC-1996; 96US-0759599.
XX 06-DEC-1995; 95US-0008233.
XX 05-DEC-1996; 96US-0759599.
XX (RABK/) RABKIN S W.
XX Krystal G, Rabkin SW;
XX WPI, 1999-394231/33.
XX Peptides that ameliorate cell death useful for treating conditions
XX associated with cellular differentiation
XX Claim 2; Column 12; 15pp; English.
XX
XX AAY25009-Y25019 are novel peptides derived from streptokinase that
XX ameliorate cell death. The products of the invention and their encoding
XX nucleic acids may be useful for treating diseases and conditions related
XX to aging, cellular differentiation, physical insult (e.g. physical
XX trauma, anoxia, hyperthermia, hypothermia, chemically induced damage,
XX and trauma to the brain, spinal cord, kidney, heart, lungs, liver, skin
XX and any other organ), viral disorders (e.g. hepatitis, retroviral
XX infections, viral encephalitis, and AIDS/HIV), neurodegenerative
XX disorders (e.g. Parkinson's disease, Alzheimer's disease, Huntington's
XX disease, cerebellar degenerations, and familial amyotrophic lateral
XX sclerosis (FALS)), cardiovascular disease (e.g. atherosclerosis,
XX myocardial infarction, heart failure, cardiomyopathy, myocardial
XX reperfusion injury, and hypertensive heart disease), immune disease (e.g.
XX rheumatoid arthritis, systemic lupus erythematosus, insulin-dependent,
XX diabetes mellitus, lupus, pernicious anaemia, dermatomyositis, erythema
XX nodosum, Sjogren's syndrome, temporal arthritis, myasthenia gravis,
XX Wegener's granulomatosis, glomerulonephritis, anti-phospholipid syndrome,
XX and autoimmune polyarthritides), a neoplastic disorder (e.g. leukemia,
XX sarcomas, myelomas, carcinomas, neuromas, melanoma, cancers of the
XX breast, brain, colon, cervix, and prostate, Hodgkin's disease and
XX non-Hodgkin's lymphoma), inflammatory disorders (e.g. inflammatory joint
XX disorders and inflammatory induced cell damage to the eye, brain and
XX other organs), ischemia or reperfusion injury (e.g. myocardial ischemia
XX and reperfusion injury, renal ischemia, spinal cord ischemia and/or
XX reperfusion injury, retinal ischemia or infarction, and stroke), toxic
XX insult (e.g. liver toxicity, pulmonary toxicity, toxic damage to other
XX organs from chemicals, radiation, and other noxious substances), macular
XX degeneration, cataract formation, pancreatitis, Crohn's disease,
XX ulcerative colitis, accelerated aging, spinal cord disease (e.g. motor
XX neuron diseases, degeneration of the spinal cord, Guillan Bare Syndrome
XX and demyelinating disease), bypass surgery, chemotherapy,
XX chemically-induced reperfusion, and therapeutics such as clozapine, AZT,
XX and anthracyclines.
XX
XX Sequence 6 AA;
XX
XX Query Match 100.0%; Score 30; DB 20; Length 6;
XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 SVDVEY 6
XX |||||
XX 1 SVDVEY 6
XX
XX RESULT 2
XX ABB80001
XX ID ABB80001 standard; peptide; 6 AA.
XX AC ABB80001;
XX XX
XX 26-JUL-2002 (first entry)
XX

DE Streptokinase derived peptide#1.
XX
XX Streptokinase, cell death; apoptosis; necrosis; nontropic;
XX neuroprotective; antiparkinsonian; anticonvulsant; cytostatic;
XX antiinflammatory; antiarthritic; antirheumatic; cardiant;
XX antiatherosclerotic; vasotropic; immunosuppressive; anti-HIV;
XX dermatological; antidiabetic; antianaemic; virucide; ophthalmological;
XX antiulcer; antibacterial; antiparasitic; neurodegenerative disease;
XX Parkinson's disease; Alzheimer's disease; Huntington's disease;
XX cerebellar degeneration; neoplastic disorder; cancer;
XX inflammatory disorder; arthritis; inflammatory joint disorders;
XX cardiovascular disease; heart failure; atherosclerosis;
XX myocardial reperfusion injury; immune disease; autoimmune disease;
XX acquired immunodeficiency syndrome; AIDS; rheumatoid arthritis;
XX systemic lupus erythematosus; diabetes mellitus; pernicious anaemia;
XX myelodegenerative diseases; viral diseases; macular degeneration;
XX cataract; Crohn's disease; ulcerative colitis; pancreatitis;
XX prion disease; aging.
XX
XX Synthetic.
XX
XX US6348567-B1.
XX
XX 19-FEB-2002.
XX
XX 19-APR-1999; 99US-0294457.
XX
XX 06-DEC-1995; 95US-008233P.
XX 05-DEC-1996; 96US-0759599.
XX
XX (MOLE-) MOLECULAR THERAPEUTICS INC.
XX
XX Krystal G, Rabkin SW;
XX WPI, 2002-266542/31.
XX
XX New peptides obtained from streptokinase, useful in ameliorating cell
XX death due to apoptosis and/or necrosis and treating neurodegenerative,
XX neoplastic, immune, cardiovascular and inflammatory disorders
XX
XX Claim 3; Column 5; 18pp; English.
XX
XX The invention relates to an isolated peptide obtained from streptokinase,
XX or its derivative or analog, which ameliorate cell death. The activity of
XX peptides of the invention may be described as, nontropic,
XX neuroprotective, antiparkinsonian, anticonvulsant, cytostatic,
XX antiinflammatory, antiarthritic, antirheumatic, cardiant,
XX antiatherosclerotic, vasotropic, immunosuppressive, anti-HIV,
XX dermatological, antidiabetic, antianaemic, virucide, ophthalmological,
XX antiulcer, antibacterial and antiparasitic. Peptides of the invention
XX ameliorates apoptosis and necrosis in a warm-blooded animal. Compositions
XX comprising peptides of the invention are useful for treating
XX neurodegenerative diseases (e.g. Parkinson's, Alzheimer's, Huntington's
XX disease and cerebellar degeneration) neoplastic disorders including
XX cancer, inflammatory disorders (e.g. arthritis, inflammatory joint
XX disorders), cardiovascular diseases (e.g. heart failure, atherosclerosis
XX and myocardial reperfusion injury), immune diseases (e.g. autoimmune
XX disease, acquired immunodeficiency syndrome (AIDS), rheumatoid arthritis,
XX systemic lupus erythematosus, diabetes mellitus, pernicious anaemia),
XX myelodegenerative diseases, viral diseases, and degenerative diseases of
XX any organ. Other disorders include macular degeneration, cataracts,
XX Crohn's disease, ulcerative colitis, cataracts, pancreatitis, infectious
XX diseases including bacteria, parasite, prion-based diseases, and
XX accelerated aging. The current sequence represents a streptokinase
XX derived peptide of the invention with an ability to ameliorate cell
XX death in cardiac myocytes.
XX
XX Sequence 6 AA;
XX
XX Query Match 100.0%; Score 30; DB 23; Length 6;
XX Best Local Similarity 100.0%; Pred. No. 7.8e+05;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
 Db 1 SVDVEY 6

RESULT 3
 ID ABB80005 standard; peptide; 16 AA.
 XX
 AC ABB80005;
 XX
 DT 26-JUL-2002 (first entry)
 XX
 DE Streptokinase derived peptide#5.
 XX
 KW Streptokinase; cell death; apoptosis; necrosis; neurotropic; neuroprotective; antiparkinsonian; anticonvulsant; cytostatic;
 KW antiinflammatory; antiarthritic; antirheumatic; cardiant;
 KW antiatherosclerotic; vasotropic; immunosuppressive; anti-HIV;
 KW dermatological; antidiabetic; antianaemic; virucide; ophthalmological;
 KW antitumor; antibacterial; antiparasitic; immunosuppressive; antirheumatic;
 KW Parkinson's disease; Alzheimer's disease; Huntington's disease;
 KW cerebellar degeneration; neoplastic disorder; cancer;
 KW inflammatory disorder; arthritis; inflammatory joint disorders;
 KW cardiovascular disease; heart failure; atherosclerosis;
 KW myocardial reperfusion injury; immune disease; autoimmune disease;
 KW acquired immunodeficiency syndrome; AIDS; rheumatoid arthritis;
 KW systemic lupus erythematosus; diabetes mellitus; pernicious anaemia;
 KW myelodegenerative diseases; viral diseases; macular degeneration;
 KW cataract; Crohn's disease; ulcerative colitis; pancreatitis;
 KW prion disease; aging.
 XX
 OS Synthetic.
 XX
 XX US6348567-B1.
 XX
 XX 19-FEB-2002.
 XX
 XX 19-APR-1999; 99US-0294457.
 XX
 XX 06-DEC-1995; 95US-008233P.
 XX
 XX 05-DEC-1996; 96US-0759599.
 XX
 XX (MOLE-) MOLECULAR THERAPEUTICS INC.
 XX
 XX Krystal G, Rabkin SW;
 XX
 XX WPI; 2002-266542/31.
 XX
 XX New peptides obtained from streptokinase, useful in ameliorating cell death due to apoptosis and/or necrosis and treating neurodegenerative, neoplastic, immune, cardiovascular and inflammatory disorders -
 XX
 XX Claim 7; Column 5; 18pp; English.
 XX
 XX The invention relates to an isolated peptide obtained from streptokinase, or its derivative or analog, which ameliorate cell death. The activity of peptides of the invention may be described as, neurotropic,
 XX neuroprotective, antiparkinsonian, anticonvulsant, cytostatic,
 XX antiinflammatory, antiarthritic, antirheumatic, cardiant,
 XX antiatherosclerotic, vasotropic, immunosuppressive, anti-HIV,
 XX dermatological, antidiabetic, antianaemic, virucide, ophthalmological, antitumor, antibacterial and antiparasitic. Peptides of the invention ameliorates apoptosis and necrosis in a warm-blooded animal. Compositions comprising peptides of the invention are useful for treating
 XX neurodegenerative diseases (e.g. Parkinson's, Alzheimer's, Huntington's disease and cerebellar degeneration) neoplastic disorders including cancer, inflammatory disorders (e.g. arthritis, inflammatory joint disorders), cardiovascular diseases (e.g. heart failure, atherosclerosis and myocardial reperfusion injury), immune diseases (e.g. autoimmune disease, acquired immunodeficiency syndrome (AIDS), rheumatoid arthritis, systemic lupus erythematosus, diabetes mellitus, pernicious anaemia), of myelodegenerative diseases, viral diseases, and degenerative diseases, of

CC any organ. Other disorders include macular degeneration, cataracts, Crohn's disease, ulcerative colitis, cataracts, pancreatitis, infectious diseases including bacteria, parasite, prion-based diseases, and accelerated aging. The current sequence represents a streptokinase derived peptide of the invention with an ability to ameliorate cell death in cardiac myocytes.

SQ Sequence 16 AA;
 Query Match 100.0%; Score 30; DB 23; Length 16;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
 Db 1 SVDVEY 6

RESULT 4
 ID AAY25013 standard; peptide; 18 AA.
 XX
 AC AAY25013;
 XX
 DT 23-AUG-1999 (first entry)
 XX
 DE Streptokinase derived peptide 5 for amelioration of cell death.
 XX
 KW Streptokinase; cell death; amelioration; treatment; disease; aging; AIDS; cellular differentiation; physical insult; trauma; anoxia; hyperthermia; hypothermia; brain; spinal cord; kidney; heart; lung; liver; skin; HIV; viral disorder; hepatitis; retroviral; infection; encephalitis; FALS;
 KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease; Huntington's disease; cerebellar degeneration; cardiovascular disease; familial amyotrophic lateral sclerosis; atherosclerosis; heart failure; infarction; heart disease; cardiomyopathy; hypertensive; myocardial; reperfusion injury; immune disease; rheumatoid arthritis; renal; retinal; systemic lupus erythematosus; insulin-dependent; diabetes mellitus; pernicious anaemia; dermatomyositis; erythema nodosum; myositis; Sjogren's syndrome; temporal arthritis; autoimmune polyarthritides; Wegener's granulomatosis; glomerulonephritis; anti-phospholipid syndrome; neoplastic disorder; leukemia; sarcoma; myeloma; carcinoma; neuroma; melanoma; cancer; breast; colon; cervix; prostate; Hodgkin's disease; non-Hodgkin's lymphoma; inflammatory disorder; stroke; ischemia; spinal cord; toxic insult; pulmonary; macular degeneration; cataract; pancreatitis; Crohn's disease; ulcerative colitis; motor neuron disease; Guillan Bare Syndrome; demyelinating disease; bypass surgery; chemotherapy; clozapine; AZT; anthracycline.
 XX
 OS Synthetic.
 XX
 XX US5917013-A.
 XX
 XX 29-JUN-1999.
 XX
 XX 05-DEC-1996; 96US-0759599.
 XX
 XX 06-DEC-1995; 95US-0008233.
 XX
 XX 05-DEC-1996; 96US-0759599.
 XX
 XX (RABK/) RABKIN S W.
 XX
 XX Krystal G, Rabkin SW;
 XX
 XX WPI; 1999-394231/33.
 XX
 XX Peptides that ameliorate cell death useful for treating conditions associated with cellular differentiation
 XX
 XX Claim 6; Column 12; 15pp; English.
 XX
 XX AAY25009-Y25019 are novel peptides derived from streptokinase that ameliorate cell death. The products of the invention and their encoding

CC nucleic acids may be useful for treating diseases and conditions related
 CC to aging, cellular differentiation, physical insult (e.g. physical
 CC trauma, anoxia, hyperthermia, hypothermia, chemically induced damage,
 CC and trauma to the brain, spinal cord, kidney, heart, lungs, liver, skin
 CC and any other organ), viral disorders (e.g. hepatitis, retroviral
 CC infections, viral encephalitis, and AIDS/HIV), neurodegenerative
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease, Huntington's
 CC disease, cerebellar degenerations, and familial amyotrophic lateral
 CC sclerosis (FALS)), cardiovascular disease (e.g. atherosclerosis,
 CC myocardial infarction, heart failure, cardiomyopathy, myocardial
 CC reperfusion injury, and hypertensive heart disease), immune disease (e.g.
 CC rheumatoid arthritis, systemic lupus erythematosus, insulin-dependent
 CC diabetes mellitus, lupus, pernicious anaemia, dermatomyositis, enythem
 CC nodosum, Sjogren's syndrome, temporal arthritis, myasthenia gravis,
 CC Wegener's granulomatosis, glomerulonephritis, anti-phospholipid syndrome,
 CC and autoimmune polyarthritides), a neoplastic disorder (e.g. leukemia,
 CC sarcomas, myelomas, carcinomas, neuromas, melanoma, cancers of the
 CC breast, brain, colon, cervix, and prostate, Hodgkin's disease and
 CC non-Hodgkin's lymphoma), inflammatory disorders (e.g. inflammatory joint
 CC disorders and inflammation induced cell damage to the eye, brain and
 CC other organs), ischemia or reperfusion injury (e.g. myocardial ischemia
 CC and reperfusion injury, renal ischemia, spinal cord ischemia and/or
 CC reperfusion injury, retinal ischemia or infarction, and stroke), toxic
 CC insult (e.g. liver toxicity, pulmonary toxicity, toxic damage to other
 CC organs from chemicals, radiation, and other noxious substances), macular
 CC degeneration, cataract formation, pancreatitis, Crohn's disease,
 CC ulcerative colitis, accelerated aging, spinal cord disease (e.g. motor
 CC neuron diseases, degeneration of the spinal cord, Guillan Bare Syndrome
 CC and demyelinating disease), bypass surgery, chemotherapy,
 CC chemically-induced reperfusion, and therapeutics such as clozapine, AZT,
 CC and anthracyclines.

XX Sequence 18 AA;

Query Match 100.0%; Score 30; DB 20; Length 18;
 Best Local Similarity 100.0%; Pred. NO. 1.8;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6

Db 1 SVDVEY 6

RESULT 5

RAY25014

ID AAY25014 standard; peptide; 20 AA.

XX AC AAY25014;

XX DT 23-AUG-1999 (first entry)

XX DE Streptokinase derived peptide 6 for amelioration of cell death.

XX Streptokinase; cell death; amelioration; treatment; disease; aging; AIDS;
 KW cellular differentiation; physical insult; trauma; anoxia; hyperthermia;
 KW hypothermia; brain; spinal cord; kidney; heart; lung; liver; skin; HIV;
 KW viral disorder; hepatitis; retroviral; infection; encephalitis; FALS;
 KW neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
 KW Huntington's disease; cerebellar degeneration; cardiovascular disease;
 KW familial amyotrophic lateral sclerosis; atherosclerosis; heart failure;
 KW infarction; heart disease; cardiomyopathy; hypertensive; myocardial;
 KW reperfusion injury; immune disease; rheumatoid arthritis; renal; retinal;
 KW systemic lupus erythematosus; insulin-dependent; diabetes mellitus;
 KW pernicious anaemia; dermatomyositis; enythem nodosum; myasthenia gravis;
 KW Sjogren's syndrome; temporal arthritis; autoimmune polyarthritides;
 KW Wegener's granulomatosis; glomerulonephritis; anti-phospholipid syndrome;
 KW neoplastic disorder; leukemia; sarcoma; myeloma; carcinoma; neuroma;
 KW melanoma; cancer; breast; colon; cervix; prostate; Hodgkin's disease;
 KW non-Hodgkin's lymphoma; inflammatory disorder; stroke; ischemia;
 KW spinal cord; toxic insult; pulmonary; macular degeneration; cataract;
 KW pancreatitis, Crohn's disease; ulcerative colitis; motor neuron disease;
 KW Guillan Bare Syndrome; demyelinating disease; bypass surgery;
 KW chemotherapy; clozapine; AZT; anthracycline.

XX Synthetic.
 OS US9917013-A.
 XX 29-JUN-1999.

XX 05-DEC-1996; 96US-0759599.

XX 06-DEC-1995; 95US-0008233.

XX 05-DEC-1996; 96US-0759599.

XX (RABK/) RABKIN S W.

XX Krystal G, Rabkin SW;

XX WPI; 1999-394231/33.

XX Peptides that ameliorate cell death useful for treating conditions
 associated with cellular differentiation

XX Claim 7; Column 12; 15pp; English.

XX AAY25009-Y25019 are novel peptides derived from streptokinase that
 ameliorate cell death. The products of the invention and their encoding
 nucleic acids may be useful for treating diseases and conditions related
 to aging, cellular differentiation, physical insult (e.g. physical
 trauma, anoxia, hyperthermia, hypothermia, chemically induced damage,
 CC and trauma to the brain, spinal cord, kidney, heart, lungs, liver, skin
 CC and any other organ), viral disorders (e.g. hepatitis, retroviral
 CC infections, viral encephalitis, and AIDS/HIV), neurodegenerative
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease, Huntington's
 CC disease, cerebellar degenerations, and familial amyotrophic lateral
 CC sclerosis (FALS)), cardiovascular disease (e.g. atherosclerosis,
 CC myocardial infarction, heart failure, cardiomyopathy, myocardial
 CC reperfusion injury, and hypertensive heart disease), immune disease (e.g.
 CC rheumatoid arthritis, systemic lupus erythematosus, insulin-dependent,
 CC diabetes mellitus, lupus, pernicious anaemia, dermatomyositis, enythem
 CC nodosum, Sjogren's syndrome, temporal arthritis, myasthenia gravis,
 CC Wegener's granulomatosis, glomerulonephritis, anti-phospholipid syndrome,
 CC and autoimmune polyarthritides), a neoplastic disorder (e.g. leukemia,
 CC sarcomas, myelomas, carcinomas, neuromas, melanoma, cancers of the
 CC breast, brain, colon, cervix, and prostate, Hodgkin's disease and
 CC non-Hodgkin's lymphoma), inflammatory disorders (e.g. inflammatory joint
 CC disorders and inflammation induced cell damage to the eye, brain and
 CC other organs), ischemia or reperfusion injury (e.g. myocardial ischemia
 CC and reperfusion injury, renal ischemia, spinal cord ischemia and/or
 CC reperfusion injury, retinal ischemia or infarction, and stroke), toxic
 CC insult (e.g. liver toxicity, pulmonary toxicity, toxic damage to other
 CC organs from chemicals, radiation, and other noxious substances), macular
 CC degeneration, cataract formation, pancreatitis, Crohn's disease,
 CC ulcerative colitis, accelerated aging, spinal cord disease (e.g. motor
 CC neuron diseases, degeneration of the spinal cord, Guillan Bare Syndrome
 CC and demyelinating disease), bypass surgery, chemotherapy,
 CC chemically-induced reperfusion, and therapeutics such as clozapine, AZT,
 CC and anthracyclines.

XX Sequence 20 AA;

Query Match 100.0%; Score 30; DB 20; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6

Db 1 SVDVEY 6

RESULT 6

ABB80006

ID ABB80006 standard; peptide; 20 AA.

XX AC ABB80006;

CC reperfusion injury, and hypertensive heart disease), immune disease (e.g. rheumatoid arthritis, systemic lupus erythematosus, insulin-dependent, diabetes mellitus, lupus, pernicious anaemia, dermatomyositis, erythema nodosum, Sjogren's syndrome, temporal arthritis, myasthenia gravis, Wegener's granulomatosis, glomerulonephritis, anti-phospholipid syndrome, CC and autoimmune polyarthritides), a neoplastic disorder (e.g. leukemia, sarcomas, myelomas, carcinomas, neuromas, melanoma, cancers of the CC breast, brain, colon, cervix, and prostate, Hodgkin's disease and non-Hodgkin's lymphoma), inflammatory disorders (e.g. inflammatory joint disorders and inflammatory induced cell damage to the eye, brain and other organs), ischemia or reperfusion injury (e.g. myocardial ischemia CC and reperfusion injury, renal ischemia, spinal cord ischemia and/or reperfusion injury, retinal ischemia or infarction, and stroke), toxic CC insult (e.g. liver toxicity, pulmonary toxicity, toxic damage to other organs from chemicals, radiation, and other noxious substances), macular CC degeneration, cataract formation, pancreatitis, Crohn's disease, ulcerative colitis, accelerated aging, spinal cord disease (e.g. motor CC neuron diseases, degeneration of the spinal cord, Guillan Bare Syndrome and demyelinating disease), bypass surgery, chemotherapy, CC chemically-induced reperfusion, and therapeutics such as clozapine, AZT, CC and anthracyclines.

XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 30; DB 20; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
| | | | |
Db 1 SVDVEY 6

RESULT 8
ID ABB80007 standard; peptide; 21 AA.
XX
AC ABB80007;
XX
DT 26-JUL-2002 (first entry)
XX
DE Streptokinase derived peptide#7.
XX
KW Streptokinase; cell death; apoptosis; necrosis; nontropic;
KW neuroprotective; antiparkinsonian; anticonvulsant; cytostatic;
KW antiinflammatory; antiarthritic; antirheumatic; cardiant;
KW antiatherosclerotic; vasotropic; immunosuppressive; anti-HIV;
KW dermatological; antidiabetic; antianaemic; virucide; ophthalmological;
KW antiulcer; antibacterial; antiparasitic; neurodegenerative disease;
KW Parkinson's disease; Alzheimer's disease; Huntington's disease;
KW cerebellar degeneration; neoplastic disorder; cancer;
KW inflammatory disorder; arthritis; inflammatory joint disorders;
KW cardiovascular disease; heart failure; atherosclerosis;
KW myocardial reperfusion injury; immune disease; autoimmune disease;
KW acquired immunodeficiency syndrome; AIDS; rheumatoid arthritis;
KW systemic lupus erythematosus; diabetes mellitus; pernicious anaemia;
KW myelodegenerative diseases; viral diseases; macular degeneration;
KW cataract; Crohn's disease; ulcerative colitis; pancreatitis;
KW prion disease; aging.
XX
OS Synthetic.
XX
PN US6348567-B1.
XX
PD 19-FEB-2002.
XX
PF 19-APR-1999; 99US-0294457.
XX
PR 06-DEC-1995; 95US-008233P.
PR 05-DEC-1996; 96US-0759599.
XX
PA (MOLE-) MOLECULAR THERAPEUTICS INC.

PI Krystal G, Rabkin SW;
XX
DR WPI; 2002-266542/31.
XX
XX New peptides obtained from streptokinase, useful in ameliorating cell death due to apoptosis and/or necrosis and treating neurodegenerative, PT neoplastic, immune, cardiovascular and inflammatory disorders -
PT
XX
PS Claim 9; Column 5; 18pp; English.
XX
XX The invention relates to an isolated peptide obtained from streptokinase, or its derivative or analog, which ameliorate cell death. The activity of CC peptides of the invention may be described as, nontropic, CC neuroprotective, antiparkinsonian, anticonvulsant, cycostatic, CC antiatherosclerotic, vasotropic, antirheumatic, cardiant, CC dermatological, antidiabetic, antianaemic, virucide, ophthalmological, CC antiulcer, antibacterial and antiparasitic. Peptides of the invention CC ameliorates apoptosis and necrosis in a warm-blooded animal. Compositions CC comprising peptides of the invention are useful for treating CC neurodegenerative diseases (e.g. Parkinson's, Alzheimer's, Huntington's disease and cerebellar degeneration) neoplastic disorders including CC cancer, inflammatory disorders (e.g. arthritis, inflammatory joint disorders), cardiovascular diseases (e.g. heart failure, atherosclerosis and myocardial reperfusion injury), immune diseases (e.g. autoimmune disease, acquired immunodeficiency syndrome (AIDS), rheumatoid arthritis, systemic lupus erythematosus, diabetes mellitus, pernicious anaemia), CC myelodegenerative diseases, viral diseases, and degenerative diseases of CC any organ. Other disorders include macular degeneration, cataracts, CC Crohn's disease, ulcerative colitis, cataracts, pancreatitis, infectious diseases including bacteria, parasite, prion-based diseases, and CC accelerated aging. The current sequence represents a streptokinase CC derived peptide of the invention with an ability to ameliorate cell CC death in cardiac myocytes.

XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 30; DB 23; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
| | | | |
Db 1 SVDVEY 6

RESULT 9
ID AAY01550 standard; peptide; 71 AA.
XX
AC AAY01550;
XX
DT 18-JUN-1999 (first entry)
XX
DE Antigenic epitope of streptokinase, spanning amino acids 138-208.
XX
XX Antigenic peptide; streptokinase; streptokinase-specific antibody;
KW thrombolytic activity; thrombolytic therapy; glomerulonephritis;
KW rheumatic fever.
XX
OS Streptococcus equisimilis.
XX
PN WO9908698-A1.
XX
PD 25-FEB-1999.
XX
PF 18-AUG-1998; 98WO-US17114.
XX
PR 18-AUG-1997; 97US-0055911.
XX
PA (GEO) GEN HOSPITAL CORP.
XX
PA (HARD) HARVARD COLLEGE.

PI Parhami-Seren B, Reed GL;
XX WPI; 1999-190113/16.
XX New polypeptides which bind streptokinase-specific antibodies -
PT useful in thrombolytic therapy
XX
XX Claim 16; Page 5; 44pp; English.
PS
CC The present sequence represents an antigenic epitope of
CC streptokinase. The specification describes a polypeptide
CC which binds to a streptokinase-specific antibody and prevents the
CC antibody binding to native streptokinase. The specification also
CC describes a synthetic polypeptide (P1) comprising an epitope which
CC binds to an streptokinase-specific antibody and reduces thrombolytic
CC activity of streptokinase. P1 is used in thrombolytic therapy, and to
CC prevent or treat glomerulonephritis and rheumatic fever.
XX
XX Sequence 71 AA;
SQ
Query Match 100.0%; Score 30; DB 20; Length 71;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVDVEY 6
Db 20 SVDVEY 25
|||
RESULT 10
AAY01558
ID AAY01558 standard; peptide; 146 AA.
XX
AC AAY01558;
XX
DT 18-JUN-1999 (first entry)
XX
DE Peptide corresponding to amino acids 148-293 of streptokinase.
XX
XX Antigenic peptide; streptokinase; streptokinase-specific antibody;
KW thrombolytic activity; thrombolytic therapy; glomerulonephritis;
KW rheumatic fever.
XX
XX Streptococcus equisimilis.
OS
XX WO9908698-A1.
XX
PD 25-FEB-1999.
XX
XX 18-AUG-1998; 98WO-US17114.
PF
XX 18-AUG-1997; 97US-0055911.
PR
XX (GEO) GEN HOSPITAL CORP.
PA (HARD) HARVARD COLLEGE.
PA
XX Parhami-Seren B, Reed GL;
PI
XX WPI; 1999-190113/16.
DR
XX New polypeptides which bind streptokinase-specific antibodies -
PT useful in thrombolytic therapy
PT
PS Disclosure; Page 13; 44pp; English.
XX
XX The present sequence represents a peptide corresponding to
CC amino acids 148-293 of streptokinase. The specification describes a
CC polypeptide which binds to a streptokinase-specific antibody and
CC prevents the antibody binding to native streptokinase. The specification
CC also describes a synthetic polypeptide (P1) comprising an epitope which
CC binds to an streptokinase-specific antibody and reduces thrombolytic
CC activity of streptokinase. P1 is used in thrombolytic therapy, and to
CC prevent or treat glomerulonephritis and rheumatic fever.
CC

XX SQ Sequence 146 AA;
Query Match 100.0%; Score 30; DB 20; Length 146;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVDVEY 6
Db 10 SVDVEY 15
|||
RESULT 11
AAY24796
ID AAY24796 standard; Protein; 150 AA.
XX
AC AAY24796;
XX
DT 26-AUG-1999 (first entry)
XX
DE Recombinant streptokinase rSKI44-293.
XX
KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW rSK; rSK; bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
XX
OS Streptococcus equisimilis.
OS
XX Synthetic.
XX
PN WO9931247-A1.
XX
PD 24-JUN-1999.
XX
PF 15-DEC-1998; 98WO-US26694.
XX
PR 15-DEC-1997; 97US-0069497.
XX
XX (HARD) HARVARD COLLEGE.
PA
XX Reed GL;
PI
XX WPI; 1999-395183/33.
DR
DR N-PSDB; AAX80494.
XX
XX N-terminally deleted streptokinase
PT
XX Claim 34; Page 67-68; 73pp; English.
XX
XX The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein;
CC a nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (1); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising
CC a bacterial fibrin-dependent plasminogen activator is useful for
CC dissolving blood clots in patients with a thrombotic condition, e.g.
CC myocardial infarction, venous thrombosis, pulmonary embolism, cerebral
CC thrombosis, graft thrombosis and arterial thrombosis. The modified
CC streptokinase can also be used in non-human mammals. Streptokinase
CC activation of plasminogen is at least 10-fold, preferably 100-fold
CC greater in the presence of fibrin than in the absence of fibrin. The
CC modified streptokinase has at least one amino acid substitution that
CC inactivates a substrate site for proteolytic cleavage. This reduces the
CC rate of degradation of the streptokinase at least two-fold. The present
CC sequence represents recombinant streptokinase, designated rSKI44-293.
XX
XX SQ Sequence 150 AA;

Query Match 100.0%; Score 30; DB 20; Length 150;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 DB 14 SVDVEY 19
 |||||

RESULT 12
 ID AAY01557 standard; peptide; 233 AA.
 XX AAY01557;
 AC AAY01557;
 DT 18-JUN-1999 (first entry)
 XX
 DE Peptide corresponding to amino acids 120-352 of streptokinase.
 XX
 KW Antigenic peptide; streptokinase; streptokinase-specific antibody;
 KW thrombolytic activity; thrombolytic therapy; glomerulonephritis;
 KW rheumatic fever.
 XX
 OS Streptococcus equisimilis.
 XX
 FN WO9908698-A1.
 XX
 PD 25-FEB-1999.
 XX
 PF 18-AUG-1998; 98WO-US17114.
 XX
 PR 18-AUG-1997; 97US-0055911.
 XX
 PA (GCHO) GEN HOSPITAL CORP.
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Parhami-Seren B, Reed GL;
 XX
 DR WPI; 1999-190113/16.
 XX
 XX New polypeptides which bind streptokinase-specific antibodies -
 PT useful in thrombolytic therapy
 PT
 PS Disclosure; Page 13; 44pp; English.
 XX
 CC The present sequence represents a peptide corresponding to
 CC amino acids 120-352 of streptokinase. The specification describes a
 CC polypeptide which binds to a streptokinase-specific antibody and
 CC prevents the antibody binding to native streptokinase. The specification
 CC also describes a synthetic polypeptide (P1) comprising an epitope which
 CC binds to an streptokinase-specific antibody and reduces thrombolytic
 CC activity of streptokinase. P1 is used in thrombolytic therapy, and to
 CC prevent or treat glomerulonephritis and rheumatic fever.
 XX
 SQ Sequence 233 AA;

Query Match 100.0%; Score 30; DB 20; Length 233;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 DB 38 SVDVEY 43
 |||||

RESULT 13
 ID AAR91599 standard; Protein; 297 AA.
 XX AAR91599;
 AC AAR91599;
 XX
 DT 08-AUG-1996 (first entry)
 XX

DE Streptokinase from pST-1.
 XX streptokinase; recombinant production; haemolytic Streptococcus.
 XX
 OS Synthetic.
 XX
 FN WO9527050-A1.
 XX
 PD 12-OCT-1995.
 XX
 PF 03-APR-1995; 95WO-CN00024.
 XX
 PR 04-APR-1994; 94CN-0112106.
 XX
 PA (UYSH-) UNIV SHANGHAI MEDICAL.
 XX
 PI Song H;
 XX
 DR WPI; 1995-358628/46.
 DR N-PSDB; AAT29961.
 XX
 PT Prepn. of streptokinase - useful for treating thrombus diseases.
 XX
 PS Example 2; Page 6-8; 24pp; Chinese.
 XX
 CC The present sequence is that of streptokinase encoded by a pST-1
 CC HindIII fragment. Recombinant streptokinase can be produced by
 CC culturing bacteria transformed with a high efficiency plasmid contg.
 CC the streptokinase gene, amplified by using haemolytic Streptococcus
 CC as the template and using inducers based on the streptokinase nucleotide
 CC sequence. The recombinant streptokinase is used to treat thrombotic
 CC diseases.
 XX
 SQ Sequence 297 AA;

Query Match 100.0%; Score 30; DB 16; Length 297;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 DB 183 SVDVEY 188
 |||||

RESULT 14
 ID AAR10196 standard; Protein; 348 AA.
 XX AAR10196;
 AC AAR10196;
 XX
 DT 28-MAR-1991 (first entry)
 XX
 DE Streptokinase (1-372, 45-68 deficient).
 XX
 KW streptokinase; thrombolytic agent; myocardial infarction.
 XX
 OS Synthetic.
 XX
 FN EP407942-A.
 XX
 PD 16-JAN-1991.
 XX
 PF 09-JUL-1990; 90EP-0113099.
 XX
 PR 11-APR-1990; 90JP-0096830.
 PR 11-JUL-1989; 89JP-0179432.
 PR 27-NOV-1989; 89JP-0307957.
 XX
 PA (SAKA) OTSUKA PHARM FACTOR.
 XX
 PI Majima E, Ogino K, Sakata Y, Uenoyama T;
 XX
 DR WPI; 1991-016179/03.

XX Synthetic gene encoding streptokinase - scale, high purity prodn.
PT of streptokinase used as a thrombolytic agent
PS Claim 7; Page 59; 76pp; English.
XX Streptokinase derivative having a C-terminal deletion (i.e. Pro
CC 373-Lys 414) and Arg 45 to Gly 68 also being deleted. Derivatives
CC show a decreased antigenicity, higher stability in blood and greater
CC selectivity of thrombolytic activity and specificity.
CC The protein is used as a thrombolytic agent in
CC patients with lung thrombus or myocardial infarction.
CC See also AAQ10230, AAR10195 and AAR10197-R10200.
XX
SQ Sequence 348 AA;
Query Match 100.0%; Score 30; DB 12; Length 348;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SVDVEY 6
| | | | |
Db 133 SVDVEY 138
| | | | |
RESULT 15
AAY24795
ID AAY24795 standard; Protein; 356 AA.
XX
AC AAY24795;
XX
DT 26-AUG-1999 (first entry)
XX
DE Recombinant streptokinase rSK59-414.
XX
KW Streptococcus; streptokinase; fibrin-dependent plasminogen activator;
KW nSK; rSK; bacterial; blood clot; thrombotic condition;
KW myocardial infarction; venous thrombosis; pulmonary embolism;
KW cerebral thrombosis; graft thrombosis; arterial thrombosis.
XX
OS Streptococcus equisimilis.
OS Synthetic.
XX
XX WO9931247-A1.
XX
XX 24-JUN-1999.
XX
XX 15-DEC-1998; 98WO-US26694.
XX
XX 15-DEC-1997; 97US-0069497.
XX
XX (HARD) HARVARD COLLEGE.
XX
XX Reed GL;
XX
XX WPI; 1999-395183/33.
XX
XX N-PSDB; AAX80493.
XX
XX N-terminally deleted streptokinase
XX
XX Claim 35; Page 65-66; 73pp; English.
XX
XX The present invention describes an isolated bacterial protein that
CC induces fibrin-dependent plasminogen activation in a pharmaceutical
CC composition for dissolving blood clots. Also described are: (1) a
CC composition comprising an isolated modified streptokinase, the
CC modification being removal of amino acid residues in the amino terminus;
CC (2) a method for dissolving a blood clot in a subject, comprising
CC administering to the subject a fibrin-dependent streptokinase protein;
CC a nucleic acid (1) encoding a modified bacterial streptokinase; (3) an
CC expression vector comprising (1); and (4) a host cell transformed with
CC the expression vector of (3). The pharmaceutical composition comprising
CC a bacterial fibrin-dependent plasminogen activator is useful for

CC dissolving blood clots in patients with a thrombotic condition, e.g.
CC myocardial infarction, venous thrombosis, pulmonary embolism, cerebral
CC thrombosis, graft thrombosis and arterial thrombosis. The modified
CC streptokinase can also be used in non-human mammals. Streptokinase
CC activation of plasminogen is at least 10-fold, preferably 100-fold.
CC greater in the presence of fibrin than in the absence of fibrin. The
CC modified streptokinase has at least one amino acid substitution that
CC inactivates a substrate site for proteolytic cleavage. This reduces the
CC rate of degradation of the streptokinase at least two-fold. The present
CC sequence represents recombinant streptokinase, designated rSK59-414.
XX
SQ Sequence 356 AA;
Query Match 100.0%; Score 30; DB 20; Length 356;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SVDVEY 6
| | | | |
Db 99 SVDVEY 104
| | | | |

Search completed: January 10, 2003, 11:46:51
Job time : 36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 11:44:11 ; Search time 10 Seconds
(without alignments)
24.886 Million cell updates/sec

Title: US-09-919-703-1
Perfect score: 30
Sequence: 1 SVDVEY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	280	1 LEC2_MEDTR	Q01807 medicago tr
2	30	100.0	440	1 STRP_STREQ	P00779 streptococc
3	30	100.0	440	1 STRP_STRPY	P10520 streptococc
4	30	100.0	440	1 STRP_STRSP	P10519 streptococc
5	29	96.7	183	1 Y34A_METJA	P81307 methanococc
6	28	93.3	311	1 DNAI1_BACSU	P06567 bacillus su
7	27	90.0	562	1 TPPI_MOUSE	O89023 mus musculu
8	27	90.0	563	1 TPPI_CANPA	Q9XSB8 canis fami
9	27	90.0	563	1 TPPI_RAT	Q9EQV6 rattus norv
10	27	90.0	774	1 STF_LAMBD	P03764 bacterioph
11	27	90.0	864	1 AGLU_MUCJA	Q92442 mucor javan
12	27	90.0	1120	1 STFR_ECOLI	P76072 escherichia
13	27	90.0	1471	1 MYSA_YEAST	P32492 saccharomyc
14	27	90.0	3344	1 POLG_FRSVH	Q01901 p genome po
15	26	86.7	47	1 YOBP_ECOLI	P76265 escherichia
16	26	86.7	148	1 YCBG_HABIN	P44161 haemophilus
17	26	86.7	205	1 RAC1_DICDI	Q9GPR2 dictyosteli
18	26	86.7	263	1 XL13_ARATH	Q91ZT5 arabidopsis
19	26	86.7	305	1 HEM3_AERPE	Q9Y9J0 aeropyrum p
20	26	86.7	307	1 RUVB_MYCEN	Q49425 mycoplasma
21	26	86.7	307	1 RUVB_MYCEN	P75242 mycoplasma
22	26	86.7	321	1 FLIM_AGRV5	Q44457 agrobacteri
23	26	86.7	326	1 VT2_MXXVL	P29825 myxoma viru
24	26	86.7	341	1 RECA_LACIA	Q59486 lactococcus
25	26	86.7	445	1 FIBG_RAT	P02680 rattus norv
26	26	86.7	463	1 SYG_MYCTU	O65932 mycobacteri
27	26	86.7	496	1 ARAA_THEMA	Q9WYB3 thermotoga
28	26	86.7	526	1 SYK_CHLTR	O84786 chlamydia t
29	26	86.7	571	1 UVRC_MYCBV	O84898 mycoplasma
30	26	86.7	611	1 ACE_HAEJE	Q10715 haematobia
31	26	86.7	615	1 ACE_DROME	Q10714 drosophila
32	26	86.7	693	1 RAPI_SCHPO	Q96T17 schizosacch
33	26	86.7	749	1 YPRA_BACSU	P50830 bacillus su

RESULT 1

ID	LEC2_MEDTR	STANDARD;	PRT;	280 AA.
AC	Q01807;			
DT	01-JUL-1993	(Rel. 26, Created)		
DT	01-JUL-1993	(Rel. 26, Last sequence update)		
DT	15-JUL-1999	(Rel. 38, Last annotation update)		
DE	Truncated lectin 2 precursor.			
GN	LEC2.			
OS	Medicago truncatula (Barrel medic)			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.			
OX	NCHI_TaxID=3880;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Jemalong;			
RX	MEDLINE=92379255; PubMed=1511126;			
RA	Bauchrowitz M.A.; Barker D.G.; Nadaud I.; Rouge P.; Lescure B.;			
RT	"Lectin genes from the legume Medicago truncatula.";			
RL	Plant Mol. Biol. 19:1011-1017(1992).			
CC	-1- MISCELLANEOUS: LEC2 IS PROBABLY NON FUNCTIONAL, SINCE A FRAMESHIFT			
CC	MUTATION LEADS TO PREMATURE TRANSLATION TERMINATION AFTER ONLY 98			
CC	AA. THE SEQUENCE BELOW IGNORES THIS FRAMESHIFT MUTATION.			
CC	-1- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION			
CC	AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE			
CC	SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.			
CC	-1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; X60387; CAA42938.1; -			
DR	HSSP; P04122; 1LOE.			
DR	InterPro; IPR000985; Lectin_lega.			
DR	InterPro; IPR001220; Lectin_legB.			
DR	Pfam; PF00138; lectin_lega; 1.			
DR	Pfam; PF00139; lectin_legB; 1.			
DR	ProDom; PD000671; Lectin_lega; 1.			
DR	ProDom; PD000711; Lectin_legB; 1.			
DR	PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.			
DR	PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.			
KW	Lectin; Manganese; Calcium; Signal; Multigene family; Glycoprotein.			
FT	SIGNAL 1 26			
FT	POTENTIAL.			
FT	CHAIN 27 280			
FT	TRUNCATED LECTIN 2.			
FT	METAL 148 148			
FT	MANGANESE (BY SIMILARITY).			
FT	METAL 150 150			
FT	MANGANESE AND CALCIUM (BY SIMILARITY).			
FT	METAL 152 152			
FT	CALCIUM (BY SIMILARITY).			
FT	METAL 154 154			
FT	CALCIUM (BY SIMILARITY).			
FT	METAL 158 158			
FT	MANGANESE AND CALCIUM (BY SIMILARITY).			
FT	METAL 170 170			
FT	MANGANESE (BY SIMILARITY).			

ALIGNMENTS

T CARBOHYD 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 T CARBOHYD 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
 Q SEQUENCE 280 AA; 30473 MW; DB68690AD8015E81 CRC64;

Query Match 100.0%; Score 30; DB 1; Length 280;
 Best Local Similarity 100.0%; Pred. No. 7.5;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 |||||

DB 208 SVDVEY 213
 |||||

RESULT 2
 ID STRP_STRQ STANDARD; PRT; 440 AA.
 AC P00779;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Streptokinase C precursor.
 GN SKC.

OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=119602;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=H46A;
 RX MEDLINE=85232082; PubMed=2989113;
 RA Malke H., Roe B., Ferretti J.J.;
 RA "Nucleotide sequence of the streptokinase gene from Streptococcus
 equisimilis H46A";
 RT Gene 34:357-362(1985).
 RN [2]

RP SEQUENCE OF 27-440.
 RL MEDLINE=83127125; PubMed=6760891;
 RJ Jackson K.W., Tang J.;
 RK "Complete amino acid sequence of streptokinase and its homology with
 serine proteases";
 RL Biochemistry 21:6620-6625(1982).

CC -!- FUNCTION: THIS PROTEIN IS NOT A PROTEASE, BUT IT ACTIVATES
 PLASMINOGEN BY COMPLEXING WITH IT. AS POTENTIAL VIRULENCE FACTOR,
 IT IS THOUGHT TO PREVENT THE FORMATION OF EFFECTIVE FIBRIN
 BARRIERS AROUND THE SITE OF INFECTION, THEREBY CONTRIBUTING TO THE
 INVASIVENESS OF THE CELLS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; X02986; AAA26974.1; -;
 CC EMBL; X72832; CAA51351.1; -;
 CC PIR; A00967; BZSO.
 CC PIR; A22801; A22801.
 CC InterPro; IPR004093; Staphylokinase.
 CC Pfam; PF02821; Staphylokinase; 3.
 CC Plasmimogen activation; Signal; Virulence.
 CC SIGNAL 1 26
 CC CHAIN 27 440 STREPTOKINASE C.
 CC VARIANT 195 195 L -> D.
 CC VARIANT 207 207 D -> L.
 CC VARIANT 300 300 EKY -> LEYK (IN REF. 2).
 CC CONFLICT 298 300 N -> D (IN REF. 2).
 CC CONFLICT 438 438 N -> D (IN REF. 2).
 CC SEQUENCE 440 AA; 50140 MW; 8FC1F22648ACCC77A CRC64;

Query Match 100.0%; Score 30; DB 1; Length 440;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 |||||

DB 208 SVDVEY 213
 |||||

RESULT 2
 ID STRP_STRQ STANDARD; PRT; 440 AA.
 AC P00779;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Streptokinase C precursor.
 GN SKC.

OS Streptococcus equisimilis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=119602;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=H46A;
 RX MEDLINE=85232082; PubMed=2989113;
 RA Malke H., Roe B., Ferretti J.J.;
 RA "Nucleotide sequence of the streptokinase gene from Streptococcus
 equisimilis H46A";
 RT Gene 34:357-362(1985).
 RN [2]

RP SEQUENCE OF 27-440.
 RL MEDLINE=83127125; PubMed=6760891;
 RJ Jackson K.W., Tang J.;
 RK "Complete amino acid sequence of streptokinase and its homology with
 serine proteases";
 RL Biochemistry 21:6620-6625(1982).

CC -!- FUNCTION: THIS PROTEIN IS NOT A PROTEASE, BUT IT ACTIVATES
 PLASMINOGEN BY COMPLEXING WITH IT. AS POTENTIAL VIRULENCE FACTOR,
 IT IS THOUGHT TO PREVENT THE FORMATION OF EFFECTIVE FIBRIN
 BARRIERS AROUND THE SITE OF INFECTION, THEREBY CONTRIBUTING TO THE
 INVASIVENESS OF THE CELLS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; X02986; AAA26974.1; -;
 CC EMBL; X72832; CAA51351.1; -;
 CC PIR; A00967; BZSO.
 CC PIR; A22801; A22801.
 CC InterPro; IPR004093; Staphylokinase.
 CC Pfam; PF02821; Staphylokinase; 3.
 CC Plasmimogen activation; Signal; Virulence.
 CC SIGNAL 1 26
 CC CHAIN 27 440 STREPTOKINASE C.
 CC VARIANT 195 195 L -> D.
 CC VARIANT 207 207 D -> L.
 CC VARIANT 300 300 EKY -> LEYK (IN REF. 2).
 CC CONFLICT 298 300 N -> D (IN REF. 2).
 CC CONFLICT 438 438 N -> D (IN REF. 2).
 CC SEQUENCE 440 AA; 50140 MW; 8FC1F22648ACCC77A CRC64;

QY 1 SVDVEY 6
 |||||

DB 183 SVDVEY 188
 |||||

RESULT 3

ID STRP_STRPY STANDARD; PRT; 440 AA.
 AC P10520;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Streptokinase A precursor.
 GN SKA OR SPY1979.

OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=SF130/13 / Serotype M1;
 RX MEDLINE=89160264; PubMed=2646590;
 RA Walter F., Siegel M., Malke H.;
 RA "Nucleotide sequence of the streptokinase gene from a Streptococcus
 pyogenes type 1 strain";
 RT Pyogenic Acids Res. 17:1261-1261(1989).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=11296296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar P.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RA "Complete genome sequence of an M1 strain of Streptococcus
 pyogenes";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).

CC -!- FUNCTION: THIS PROTEIN IS NOT A PROTEASE, BUT IT ACTIVATES
 PLASMINOGEN BY COMPLEXING WITH IT. AS POTENTIAL VIRULENCE FACTOR,
 IT IS THOUGHT TO PREVENT THE FORMATION OF EFFECTIVE FIBRIN
 BARRIERS AROUND THE SITE OF INFECTION, THEREBY CONTRIBUTING TO THE
 INVASIVENESS OF THE CELLS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; X13399; CAA31765.1; -;
 CC EMBL; AE006620; AAK34665.1; -;
 CC PIR; S02724; S02724.
 CC InterPro; IPR004093; Staphylokinase.
 CC Pfam; PF02821; Staphylokinase; 3.
 CC Plasmimogen activation; Signal; Virulence; Complete proteome.
 CC SIGNAL 1 26
 CC CHAIN 27 440 STREPTOKINASE A.
 CC CONFLICT 163 163 L -> V (IN REF. 1).
 CC CONFLICT 345 345 R -> G (IN REF. 1).
 CC CONFLICT 373 373 D -> N (IN REF. 1).
 CC CONFLICT 428 428 D -> Y (IN REF. 1).
 CC CONFLICT 438 438 K -> N (IN REF. 1).
 CC SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;

Query Match 100.0%; Score 30; DB 1; Length 440;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 |||||

DB 183 SVDVEY 188
 |||||

RESULT 3
 ID STRP_STRPY STANDARD; PRT; 440 AA.
 AC P10520;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Streptokinase A precursor.
 GN SKA OR SPY1979.

OS Streptococcus pyogenes.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1314;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=SF130/13 / Serotype M1;
 RX MEDLINE=89160264; PubMed=2646590;
 RA Walter F., Siegel M., Malke H.;
 RA "Nucleotide sequence of the streptokinase gene from a Streptococcus
 pyogenes type 1 strain";
 RT Pyogenic Acids Res. 17:1261-1261(1989).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
 RX MEDLINE=21192684; PubMed=11296296;
 RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
 RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
 RA Qian Y., Jia H.G., Najjar P.Z., Ren Q., Zhu H., Song L., White J.,
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
 RA "Complete genome sequence of an M1 strain of Streptococcus
 pyogenes";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).

CC -!- FUNCTION: THIS PROTEIN IS NOT A PROTEASE, BUT IT ACTIVATES
 PLASMINOGEN BY COMPLEXING WITH IT. AS POTENTIAL VIRULENCE FACTOR,
 IT IS THOUGHT TO PREVENT THE FORMATION OF EFFECTIVE FIBRIN
 BARRIERS AROUND THE SITE OF INFECTION, THEREBY CONTRIBUTING TO THE
 INVASIVENESS OF THE CELLS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; X13399; CAA31765.1; -;
 CC EMBL; AE006620; AAK34665.1; -;
 CC PIR; S02724; S02724.
 CC InterPro; IPR004093; Staphylokinase.
 CC Pfam; PF02821; Staphylokinase; 3.
 CC Plasmimogen activation; Signal; Virulence; Complete proteome.
 CC SIGNAL 1 26
 CC CHAIN 27 440 STREPTOKINASE A.
 CC CONFLICT 163 163 L -> V (IN REF. 1).
 CC CONFLICT 345 345 R -> G (IN REF. 1).
 CC CONFLICT 373 373 D -> N (IN REF. 1).
 CC CONFLICT 428 428 D -> Y (IN REF. 1).
 CC CONFLICT 438 438 K -> N (IN REF. 1).
 CC SEQUENCE 440 AA; 49924 MW; D6227EF040B758DB CRC64;

```

Db 183 SVDVEY 188

RESULT 4
STRP_STRSP STANDARD; PRT; 440 AA.
AC P10519;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Streptokinase G precursor.
GN SKG.
OS Streptococcus sp. (strain 19099).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89160265; PubMed=2922269;
RA Walter F., Siegel M., Malke H.;
RT "Nucleotide sequence of the streptokinase gene from a group-G
RT Streptococcus.";
RL Nucleic Acids Res. 17:1262-1262(1989).
CC -1- FUNCTION: THIS PROTEIN IS NOT A PROTEASE, BUT IT ACTIVATES
CC PLASMINOGEN BY COMPLEXING WITH IT. AS POTENTIAL VIRULENCE FACTOR,
CC IT IS THOUGHT TO PREVENT THE FORMATION OF EFFECTIVE FIBRIN
CC BARRIERS AROUND THE SITE OF INFECTION, THEREBY CONTRIBUTING TO THE
CC INVASIVENESS OF THE CELLS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X13400; CAA31766.1; -
DR PIR; S02723; S02723.
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
KW Plasminogen activation; Signal; Virulence.
FT SIGNAL 1 26
FT CHAIN 27 440 STREPTOKINASE G.
SQ SEQUENCE 440 AA; 50199 MW; 5521F8825FE1B6EA CRC64;

Query Match 100.0%; Score 30; DB 1; Length 440;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
Db 183 SVDVEY.188

RESULT 5
Y34A_METJA STANDARD; PRT; 183 AA.
ID Y34A_METJA
AC P81307;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0347.1.
GN MJ0347.1.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,

Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Fraser C.W., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67488; AAB98344.1; -
DR TIGR; MJ0347.1; -
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 183 AA; 21756 MW; 769FCF6808E86C895 CRC64;

Query Match 96.7%; Score 29; DB 1; Length 183;
Best Local Similarity 83.3%; Pred. No. 8.2;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
Db 11 SVDIEY 16

RESULT 6
DNAL_BACSU STANDARD; PRT; 311 AA.
ID DNAL_BACSU
AC P06567;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Primosomal protein dnal.
GN DNAL.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87117549; PubMed=3027671;
RA Ogasawara N., Moriya S., Mazza P.G., Yoshikawa H.;
RT "Nucleotide sequence and organization of dnaB gene and neighbouring
RT genes on the Bacillus subtilis chromosome.";
RL Nucleic Acids Res. 14:9989-9999(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=97124191; PubMed=8969504;
RA Wipat A., Carter N., Brignell C.S., Guy J.B., Piper K.,
RA Sanders J., Emmerson P.T., Harwood C.R.;
RT "The dnaB-pheA (256 degrees-240 degrees) region of the Bacillus
RT subtilis chromosome containing genes responsible for stress
RT responses, the utilization of plant cell walls and primary
RT metabolism.";
RL Microbiology 142:3067-3078(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes
RT in the 200 kb rnb-dnaB region.";
RL Microbiology 143:3431-3441(1997).
RN [4]
RP SEQUENCE FROM N.A.

```

RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Bortles R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschini C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot P., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferraro E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi E., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golligly E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.P., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Takeuchi M., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis".
RL Nature 390:249-256 (1997).
RN [5]
RP SEQUENCE OF 1-206 FROM N.A.
RX MEDLINE=87118226; PubMed=3027697;
RA Hoshino T., McKenzie T., Schmidt S., Tanaka T., Sueoka N.;
RT "Nucleotide sequence of Bacillus subtilis dnaB: a gene essential for
RT DNA replication initiation and membrane attachment.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:653-657 (1987).
RN [6]
RP SIMILARITY TO DNAA.
RX MEDLINE=92195821; PubMed=1549481;
RA Koonin E.V.;
RT "Archaeobacterial virus SSV1 encodes a putative DnaA-like protein.";
RL Nucleic Acids Res. 20:1143-1143 (1992).
RN [7]
RP IDENTIFICATION.
RX MEDLINE=95291463; PubMed=7773414;
RA Bruand C., Ehrlich S.D.;
RT "The Bacillus subtilis dnaI gene is part of the dnaB operon.";
RL Microbiology 141:1199-1200 (1995).
CC -|- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -|- SIMILARITY: SOME, TO DNAA FROM VARIOUS BACTERIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; X04963; CAA28633.1; -
DR ENBL; 275208; CAA99605.1; -
DR ENBL; AF008220; AAC00359.1; -
DR ENBL; Z99118; CAB14858.1; -
DR ENBL; M5183; AAA22405.1; -
DR PIR; B24720; IQB544.
DR PIR; C26580; C26580.
DR Subtilist; BG10359; dnaI.
KW Primosome; DNA replication; ATP-binding; Complete proteome.
FT NP_BIND 168 175

FT CONFLICT 19 19 K -> N (IN REP. 5).
FT CONFLICT 24 24 M -> T (IN REP. 5).
SQ SEQUENCE 311 AA; 36114 MW; A86FC9A86841264 CRC64;
Query Match 93.3%; Score 28; DB 1; Length 311;
Best Local Similarity 66.7%; Pred. No. 25;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVDVEY 6
DB 93 SIDIEY 98
RESULT 7
TPPI_MOUSE STANDARD; PRT; 562 AA.
AC Q89023; Q90Q57;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tripeptidyl-peptidase I precursor (EC 3.4.14.9) (TPP-1) (Tripeptidyl
DE aminopeptidase) (lysosomal pepstatin insensitive protease) (LPIC).
GN CLN2 OR TPPI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99142696; PubMed=9989590;
RA Vines D.J., Warburton M.J.;
RT "Classical late infantile neuronal ceroid lipofuscinosis fibroblasts
RT are deficient in lysosomal tripeptidyl peptidase I.";
RL FEBS Lett. 443:131-135 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20028316; PubMed=10556422;
RA Katz M.L., Liu P.-C., Grob-Nunn S.E., Shibuya H., Johnson G.S.;
RT "Characterization and chromosomal mapping of a mouse ortholog of the
RT late-infantile ceroid-lipofuscinosis gene CLN2.";
RL Mamm. Genome 10:1050-1053 (1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Sleat D.E., Lobel P.;
RT "Murine homologue of the lysosomal pepstatin insensitive protease
RT which is deficient in human classical late infantile neuronal ceroid
RT lipofuscinosis.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yanakura I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Offelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
CC -|- FUNCTION: Lysosomal serine protease with tripeptidyl-peptidase I

activity. May act as a non-specific lysosomal peptidase which generates tripeptides from the breakdown products produced by lysosomal proteinases. Requires substrates with an unsubstituted N-terminus (By similarity).

-1- CATALYTIC ACTIVITY: Release of an N-terminal tripeptide from a polypeptide, but also endopeptidase activity.

-1- SUBCELLULAR LOCATION: Lysosomal.

-1- PTM: Activated by autocatalytic proteolytical processing upon acidification (By similarity).

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S53.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AJ011912; CAA09863.1; ALT_INIT.
 EMBL; AF124599; AAD32573.1; -
 EMBL; AF111172; AAD03083.1; -
 EMBL; AK002418; BAB22085.1; -
 MEROPS; S53.003; -
 MGD; MG1:1336194; Cln2.
 Glycoprotein.

FT SIGNAL 1 19 BY SIMILARITY.
 FT PROPEP 20 194 REMOVED IN MATURE FORM (BY SIMILARITY).
 FT CHAIN 195 562 TRIPEPTIDYL-PEPTIDASE I.
 FT ACT_SITE 359 562 BY SIMILARITY.
 FT ACT_SITE 474 474 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 516 516 BY SIMILARITY.
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 1 1 M -> V (IN REF. 1).
 FT CONFLICT 562 562 P -> LPPFVP (IN REF. 1).
 SQ SEQUENCE 562 AA; 61342 MW; 0AF8163EAL66396 CRC64;

Query Match 90.0%; Score 27; DB 1; Length 562;
 Best Local Similarity 83.3%; Pred. No. 83;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 | : | | |
 Db 273 SLDVEY 278

RESULT 8
 TPPI_CANFA
 ID TPPI_CANFA STANDARD; PRT; 563 AA.
 AC Q9XSB6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tripeptidyl-peptidase I precursor (EC 3.4.14.9) (TPP-I) (Tripeptidyl
 DE aminopeptidase) (Lysosomal pepstatin insensitive protease) (LPIIC).
 GN CLN2 OR TPPI.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Pissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liu P.-C., Katz M.L., Siakotos A.N., Grob S.E., Johnson G.S.;
 RA "Coding sequence and exon/intron organization of the canine CLN2 gene
 RT and its exclusion as the locus for ceroid lipofuscinosis in English
 RT setter dogs";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Lysosomal serine protease with tripeptidyl-peptidase I

activity. May act as a non-specific lysosomal peptidase which generates tripeptides from the breakdown products produced by lysosomal proteinases. Requires substrates with an unsubstituted N-terminus (By similarity).

-1- CATALYTIC ACTIVITY: Release of an N-terminal tripeptide from a polypeptide.

-1- SUBCELLULAR LOCATION: Lysosomal.

-1- PTM: Activated by autocatalytic proteolytical processing upon acidification (By similarity).

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S53.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF114167; AAD25043.1; -
 MEROPS; S53.003; -
 KW Hydrolyase; Protease; Serine protease; Zymogen; Signal; Lysosome;
 KW Glycoprotein.

FT SIGNAL 1 19 BY SIMILARITY.
 FT PROPEP 20 195 REMOVED IN MATURE FORM (BY SIMILARITY).
 FT CHAIN 196 563 TRIPEPTIDYL-PEPTIDASE I.
 FT ACT_SITE 360 360 BY SIMILARITY.
 FT ACT_SITE 475 475 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 517 517 BY SIMILARITY.
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 563 AA; 61362 MW; 21465A4C34934F4 CRC64;

Query Match 90.0%; Score 27; DB 1; Length 563;
 Best Local Similarity 83.3%; Pred. No. 83;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
 | : | | |
 Db 274 SLDVEY 279

RESULT 9
 TPPI_RAT
 ID TPPI_RAT STANDARD; PRT; 563 AA.
 AC Q9EQV6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tripeptidyl-peptidase I precursor (EC 3.4.14.9) (TPP-I) (Tripeptidyl
 DE aminopeptidase) (Lysosomal pepstatin insensitive protease) (LPIIC).
 GN CLN2 OR TPPI.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Du P., Kato S., Li Y., Maeda T., Yamane T., Yamamoto S., Fujiwara M.,
 RA Yamamoto Y., Nishi K., Ohkubo I.;
 RT "Rat tripeptidyl peptidase I: its purification and molecular
 RT cloning";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 196-217; 374-392 AND 395-429, AND CHARACTERIZATION.
 RC TISSUE=Spleen;
 RX MEDLINE=98323562; PubMed=9659384;
 RA Vines D.J., Warburton M.J.;
 RT "Purification and characterisation of a tripeptidyl aminopeptidase I

RT from rat spleen.";
 RL Biochim. Biophys. Acta 1384:233-242(1998).
 CC -!- FUNCTION: Lysosomal serine protease with tripeptidyl-peptidase I
 CC activity. May act as a non-specific lysosomal peptidase which
 CC generates tripeptides from the breakdown products produced by
 CC lysosomal proteinases. Requires substrates with an unsubstituted
 CC N-terminus. Maximum activity at pH 4, unstable above pH 7.
 CC -!- CATALYTIC ACTIVITY: Release of an N-terminal tripeptide from a
 CC polypeptide.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- PTM: Activated by autocatalytic proteolytical processing upon
 CC acidification (By similarity).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S53.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AB043870; BAB18570.1;
 KW Hydrolase; Protease; Serine protease; Zymogen; Signal; Lysosome;
 KW Glycoprotein.
 FT SIGNAL 1 19 BY SIMILARITY
 FT PROPEP 20 195 REMOVED IN NATURE FORM.
 FT CHAIN 196 563 TRIPEPTIDYL-PEPTIDASE I.
 FT ACT_SITE 360 360 BY SIMILARITY.
 FT ACT_SITE 475 475 NUCLEOPHILE (BY SIMILARITY).
 FT ACT_SITE 517 517 BY SIMILARITY.
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 210 210 N -> A (IN REF. 2).
 FT CONFLICT 216 217 VG -> SQ (IN REF. 2).
 FT CONFLICT 389 391 GGT -> SPP (IN REF. 2).
 SQ SEQUENCE 563 AA; 61332 MW; B54F3C86205DFEC1 CRC64;
 Query Match 90.0%; Score 27; DB 1; Length 563;
 Best Local Similarity 83.3%; Pred. No. 83;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVDVEY 6
 Db 274 SLDVEY 279
 RESULT 10
 ID STFL LAMB STANDARD; PRT; 774 AA.
 AC P03764; P03745;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Side tail fiber protein.
 GN STFL.
 OS Bacteriophage lambda.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
 OC Lambda-like viruses.
 OX NCBI_TaxID=10710;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83189071; PubMed=6221115;
 RA Sangher F., Coulson A.R., Hong G.F., Hill D.F., Petersen G.B.;
 RT "Nucleotide sequence of bacteriophage lambda DNA."
 RL J. Mol. Biol. 162:729-773(1982).
 CC [2]
 CC IDENTIFICATION AS STFL.
 RP MEDLINE=92165720; PubMed=1531649;
 RX Hagggaard-Ljungquist E., Halling C., Calendar R.;

RT "DNA sequences of the tail fiber genes of bacteriophage P2: evidence
 RT for horizontal transfer of tail fiber genes among unrelated
 RT bacteriophages.";
 RL J. Bacteriol. 174:1462-1477(1992).
 RN [3]
 RP RECONSTRUCTION.
 RX MEDLINE=93068310; PubMed=1439823;
 RA Hendrix R.W., Duda R.L.;
 RT Bacteriophage lambda PaPa: not the mother of all lambda phages.";
 RL Science 258:1145-1148(1992).
 CC -!- MISCELLANEOUS: The common laboratory strain of bacteriophage
 CC lambda; lambda PaPa; carries a frameshift mutation relative to Ur-
 CC lambda, the original isolate. The Ur-lambda virions have thin,
 CC jointed tail fibers (side tail fibers) that are absent from lambda
 CC wild type. Relative to lambda PaPa, Ur-lambda has expanded
 CC receptor specificity and adsorbs to E.coli cells more rapidly.
 CC -!- SIMILARITY: BELONGS TO THE TAIL FIBER FAMILY.
 CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT WAS
 CC CORRECTED IN POSITION TO 396 TO RECREATE THE ORIGINAL STP PRTOEIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; J02459; AAA96555.1; ALT_FRAME.
 DR EMBL; J02459; AAA96557.1; ALT_FRAME.
 DR PIR; A04389; OXBPII.
 DR PIR; A04370; OXBPII.
 DR InterPro; IPR004089; Chmtaxis trans.
 DR InterPro; IPR005003; Phage_fiber.
 DR Pfam; PF03335; Phage_fiber; 6.
 KW Fiber protein.
 SQ SEQUENCE 774 AA; 77527 MW; CDD1DP85E919123B CRC64;
 Query Match 90.0%; Score 27; DB 1; Length 774;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVDVEY 6
 Db 50 SMDVEY 55
 RESULT 11
 ID AGLU MUCJA STANDARD; PRT; 864 AA.
 AC Q92442;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Alpha-glucosidase precursor (EC 3.2.1.20) (Maltase).
 OS Mucor javanicus
 OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
 OC Mucor.
 OX NCBI_TaxID=51122;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=IFO 4570;
 RX MEDLINE=96271012; PubMed=8930045;
 RA Sugimoto M., Suzuki Y.;
 RT "Molecular cloning, sequencing, and expression of a cDNA encoding
 RT alpha-glucosidase from Mucor javanicus."
 RL J. Biochem. 119:500-505(1996).
 CC -!- FUNCTION: HYDROLYSES NOT ONLY MALTO-OLIGOSACCHARIDES BUT ALSO
 CC SOLUBLE STARCH.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-
 CC linked D-glucose residues with release of D-glucose.
 CC -!- SIMILARITY: BELONGS TO FAMILY 31 OF GLYCOSYL HYDROLASES.
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR ENBL; D67034; BAA11053.1; -
 DR InterPro; IPR000322; Glyco_hydro_31.
 DR Pfam; PF01055; Glyco_hydro_31; 1.
 DR PROSITE; PS00129; GLYCOSYL_HYDROL_F31_1; 1.
 DR PROSITE; PS00707; GLYCOSYL_HYDROL_F31_2; FALSE_NEG.
 KW Hydrolase; Glycosidase; Glycoprotein; Signal.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 864 ALPHA-GLUCOSIDASE.
 FT ACT_SITE 430 430 BY SIMILARITY.
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 500 500 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 734 734 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 864 AA; 98761 MW; CFAB4759DC431403 CRC64;

Query Match 90.0%; Score 27; DB 1; Length 864;
 Best Local Similarity 83.3%; Pred. No. 1.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVVEY 6
 :|||||
 DB 75 TVDVEY 80

RESULT 12

STPR_ECOLI
 ID STPR_ECOLI STANDARD; PRT; 1120 AA.
 AC P76072; P77560;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Side tail fiber protein homolog from lambdaoid prophage Rac.
 GN STPR OR B1372.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Itoh T.,
 RA Kasai H., Kashimoto K., Kimura S., Kitagawa M., Kitagawa M.,
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N.,
 RA Sampel G., Seki Y., Sivasubram S., Tagami H., Takeda J.,
 RA Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;
 RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 28.0-40.1 min region on the linkage map";
 RL DNA Res. 3:363-377 (1996).
 CC -1- SIMILARITY: BELONGS TO THE TAIL FIBER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC
 DR ENBL; AB000234; AAC74454.1; ALT_INIT.
 DR ENBL; D90774; BAA14966.1; -
 DR ENBL; D90775; BAA14975.1; -
 DR EcoGene; EGI3370; stfr.
 DR InterPro; IPR004089; Chmtaxis transd.
 DR InterPro; IPR005003; Phage_fiber.
 DR InterPro; IPR005068; Phage_fiber_2.
 DR Pfam; PF03335; Phage_fiber; 6.
 DR Pfam; PF03406; Phage_fiber_2; 1.
 KW Hypothetical protein; Fiber protein; Repeat; Complete proteome.
 SQ SEQUENCE 1120 AA; 113779 MW; 542E9D71EE795B4 CRC64;

Query Match 90.0%; Score 27; DB 1; Length 1120;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVVEY 6
 :|||||
 DB 50 SMDVEY 55

RESULT 13

MYSA_YEAST
 ID MYSA_YEAST STANDARD; PRT; 1471 AA.
 AC P32492;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin-4 isoform.
 GN MYO4 OR SHE1 OR YAL029C OR FUN22.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94334370; PubMed=8056830;
 RA Haarer B.K., Petzold A., Lillie S.H., Brown S.S.;
 RT "Identification of MYO4, a second class V myosin gene in yeast.";
 RL J. Cell Sci. 107:1055-1064 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=G288c / AB972;
 RX MEDLINE=95249563; PubMed=7731988;
 RA Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,
 RA Hall J., Ouellette B.F.P., Keng T., Barton A.B., Su Y., Davies C.K.,
 RA Storms R.K.;
 RT "The nucleotide sequence of chromosome I from Saccharomyces
 cerevisiae";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813 (1995).
 CC -1- FUNCTION: SEEMS TO BE RESPONSIBLE FOR THE ACCUMULATION IN DAUGHTER
 CC CELLS OF ASH1, A REPRESSOR OF THE HO ENDONUCLEASE.
 CC -1- SUBCELLULAR LOCATION: ACCUMULATES PREFERENTIALLY IN GROWING BUDS.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 3 IQ DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M90057; AAC37409.1; --
DR EMBL; U12980; AAC05003.1; --
DR PIR; S30790; S30790.
DR HSSP; P10587; 1BR2.
DR SGD; S0000027; MYO4.
DR InterPro; IPR002710; DIL.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF00612; IQ; 5.
DR Pfam; PF01843; DIL; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000335; myosin_head; 1.
DR ProDom; PD003376; DIL; 1.
DR SMART; SM00015; IQ; 2.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 2.
KW Myosin; ATP-binding; Actin-binding; Calmodulin-binding; Repeat;
KW Coiled coil; Alkylation.
FT DOMAIN 1 777 MYOSIN HEAD-LIKE.
FT DOMAIN 876 898 IQ 1.
FT DOMAIN 899 928 IQ 2.
FT DOMAIN 938 1063 COILED COIL.
FT DOMAIN 1064 1471 CARBOXYL-TERMINAL.
FT NP BIND 165 172 ATP (POTENTIAL).
FT MOD RES 688 688 ALKYLATION (BY SIMILARITY).
FT MOD RES 698 698 ALKYLATION (BY SIMILARITY).
SQ SEQUENCE 1471 AA; 169343 MW; E79C0FE72B041E95 CRC64;
Query Match 90.0%; Score 27; DB 1; Length 1471;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SVDVEY 6
Db 568 ADVDEY 573
RESULT 14
POLG_PRSVH STANDARD; PRT; 3344 AA.
AC Q01901;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: N-terminal protein (P1); Helper
DE component proteinase (EC 3.4.22.45) (HC-Pro); Protein P3; 6 kDa
DE protein 1 (6k1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2
DE (6k2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)
DE (NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear
DE inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase) (EC
DE 2.7.7.48); Coat protein (CP)].
OS Papaya ringspot virus (strain P / mutant HA).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=31731;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang C.H., Bau H.J., Yeh S.D.;
RT "Comparison of the nuclear inclusion b protein and coat protein genes
RT of five papaya ringspot virus strains distinct in geographic origin
RT and pathogenicity."
RL Phytopathology 84:1205-1210(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93019006; PubMed=1402799;
RA Yeh S.D., Jan F.J., Chiang C.H., Doong T.J., Chen M.C.,
RA Chung P.H., Bau H.J.;
RT "Complete nucleotide sequence and genetic organization of papaya
RT ringspot virus RNA."
RL J. Gen. Virol. 73:2531-2541(1992).

[3]
RN SEQUENCE OP 2561-3344 FROM N.A.
RP MEDLINE=93090098; PubMed=1456896;
RX Wang C.H., Yeh S.D.;
RT "Nucleotide sequence comparison of the 3'-terminal regions of severe,
RT mild, and non-papaya infecting strains of papaya ringspot virus.";
RL Arch. Virol. 127:345-354(1992)
CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHD
CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
CC MAY BE INVOLVED IN REPLICATION.
CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -1- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
CC further restricted by preferences for the amino acids in P6 - P1,
CC that vary with the species of potyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
CC Gln+ (Ser or Gly) for the enzyme from tobacco etch virus. The
CC natural substrate is the viral polyprotein, but other proteins and
CC oligopeptides containing the appropriate consensus sequence are
CC also cleaved.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -1- CATALYTIC ACTIVITY: Hydrolyzes a Gly-|-Gly bond at its own C-
CC terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-|-Gly, in the
CC processing of the potyviral polyprotein.
CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
CC POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC INDIVIDUAL PROTEINS.
CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPROTEIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; X67673; CAA47905.1; --
CC EMBL; S46722; AAB3789.1; --
CC EMBL; X67672; CAA47904.1; --
CC PIR; S24785; S24785.
CC MEROPS; C06.009; --
CC MEROPS; C06.001; --
CC MEROPS; S30.001; --
CC InterPro; IPR001410; DEAD.
CC InterPro; IPR001650; Helicase_C.
CC InterPro; IPR001730; Peptidase_C4.
CC InterPro; IPR001456; Peptidase_C6.
CC InterPro; IPR002540; Poty_P1.
CC InterPro; IPR001592; Poty_coat.
CC InterPro; IPR001205; RNA_Pol_P3D.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00271; helicase_C; 1.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00767; Poty_coat; 1.
CC Pfam; PF00851; Peptidase_C6; 1.
CC Pfam; PF00863; Peptidase_C4; 1.
CC Pfam; PF01577; Poty_P1; 1.
CC PRINTS; PR00966; NIAPOTYPTASE.
CC SMART; SM00487; DEXDC; 1.
CC SMART; SM00490; HELICc; 1.
KW Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KW ATP-binding.
FT CHAIN 1 529 N-TERMINAL PROTEIN (BY SIMILARITY).
FT CHAIN 530 1149 HELPER COMPONENT PROTEINASE
FT (BY SIMILARITY).
FT CHAIN 1150 ? PROTEIN P3 (BY SIMILARITY).
FT CHAIN ? 1401 6 KDA PROTEIN 1 (BY SIMILARITY).

```

FT CHAIN 1402 2036 CYTOPLASMIC INCLUSION PROTEIN (BY
FT CHAIN 2037 2093 SIMILARITY).
FT CHAIN 2094 ? 6 KDA PROTEIN 2 (BY SIMILARITY).
FT CHAIN ? 2520 GENOME-LINKED PROTEIN (BY SIMILARITY).
FT CHAIN ? 2520 NUCLEAR INCLUSION PROTEIN A
FT CHAIN 2521 3037 (BY SIMILARITY).
FT CHAIN 2521 3037 NUCLEAR INCLUSION PROTEIN B
FT CHAIN 3038 3344 (BY SIMILARITY).
FT CHAIN 2156 2156 COAT PROTEIN (BY SIMILARITY).
FT BINDING 2156 2156 COVALENT LINKAGE OF VIRAL RNA (BY
FT NP BIND 1486 1493 SIMILARITY).
FT SEQUENCE 3344 AA; 381040 MW; E90CD7523AC5243D CRC64;

Query Match 90.0%; Score 27; DB 1; Length 3344;
Best Local Similarity 83.3%; Pred. No. 5.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6
Db 149 SVDLEY 154

RESULT 15
YOFB ECOLI
ID YOFB ECOLI STANDARD; PRT; 47 AA.
AC P76265;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein yofB.
GN YOFB OR B1824 OR Z2869 OR ECS2534.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000276; AAC74894.1; -
CC EMBL; AE005405; AAG56813.1; -
CC EMBL; AP002558; BAB35957.1; -
CC EcoGene; EGI4382; yofB.
CC Hypothetical protein; Complete proteome.
CC SEQUENCE 47 AA; 5212 MW; 54E439F3531D708F CRC64;

Query Match 86.7%; Score 26; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDVVEY 6
Db 14 VDVVEY 18

Search completed: January 10, 2003, 11:46:09
Job time : 11 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 11:45:36 ; Search time 15 Seconds
(without alignments)
38.454 Million cell updates/sec

Title: US-09-919-703-1

Perfect score: 30

Sequence: 1 SVDVEY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	100.0	128	2 S77671	streptokinase A (B
2	30	100.0	157	2 AH3460	arsenate reductase
3	30	100.0	279	2 T09620	probable lectin 2
4	30	100.0	415	1 BZSO	streptokinase (EC
5	30	100.0	440	1 A22801	streptokinase G pr
6	30	100.0	440	2 S02723	streptokinase A pr
7	30	100.0	440	2 S02724	streptokinase G pr
8	30	100.0	1274	2 T10729	transferrin-like p
9	29	96.7	186	2 T31951	hypothetical prote
10	29	96.7	215	2 C97297	probable phosphata
11	29	96.7	1207	2 T16011	hypothetical prote
12	29	96.7	1595	2 T31082	endo-1,4-beta-xyla
13	28	93.3	311	1 IOBS44	primosome componen
14	28	93.3	629	2 T27619	hypothetical prote
15	27	90.0	50	2 A60745	major outer membra
16	27	90.0	180	2 C88465	protein B0244.9 [I
17	27	90.0	271	2 G90898	probable tail fibe
18	27	90.0	271	2 B85942	hypothetical prote
19	27	90.0	368	2 JN0848	outer membrane 40K
20	27	90.0	368	2 S46435	porin precursor -
21	27	90.0	371	2 S46436	porin foma precurs
22	27	90.0	375	2 G85631	hypothetical prote
23	27	90.0	391	2 G81704	probable transcrip
24	27	90.0	399	2 E70598	probable PPG prote
25	27	90.0	401	1 OXBp1L	hypothetical prote
26	27	90.0	407	2 G90907	probable tail fibe
27	27	90.0	437	2 E90368	probable tail fibe
28	27	90.0	437	2 E90996	probable tail fibe
29	27	90.0	437	2 H90854	probable tail fibe

30 27 90.0 438 2 D90734 probable tail fibe
31 27 90.0 439 2 C90769 probable tail fibe
32 27 90.0 439 2 R85816 probable tail fibe
33 27 90.0 439 2 A85719 probable tail fibe
34 27 90.0 439 2 A85741 hypothetical prote
35 27 90.0 440 2 R85584 probable tail comp
36 27 90.0 621 2 H81971 probable ABC trans
37 27 90.0 621 2 C81026 ABC transporter, A
38 27 90.0 831 2 E96830 hypothetical prote
39 27 90.0 864 2 Jc4624 alpha-glucosidase
40 27 90.0 971 2 B90835 probable tail fibe
41 27 90.0 973 2 C85693 probable membrane
42 27 90.0 1077 2 D90387 peptidase related
43 27 90.0 1122 2 G84887 probable tail fibe
44 27 90.0 1471 1 S30790 myosin MYO4 - years
45 27 90.0 3344 2 JQ1899 genome polyprotein

ALIGNMENTS

RESULT 1

S77671
Streptokinase A (EC 3.4.-.-) (alleles 2 and 3) - Streptococcus pyogenes (fragment)
C:Species: Streptococcus pyogenes
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 22-Jun-1999
C:Accession: S77671; S77672
R:Kapur, V.; Kanjilal, S.; Hamrick, M.R.; Li, L.L.; Whittam, T.S.; Sawyer, S.A.; Musser
Mol. Microbiol. 16, 509-519, 1995
A>Title: Molecular population genetic analysis of the streptokinase gene of Streptococ
A:Reference number: S77671; MUID:96037795; PMID:7565111
A:Accession: S77671
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <AP>
A:Cross-references: EMBL:U25853; NID:9818908; PIDN:AAA85729.1; PID:9818909
A:Experimental source: strain ET1/M1
A>Note: allele 2
A:Accession: S77672
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-128 <AA2>
A:Cross-references: EMBL:U25854; NID:9818910; PIDN:AAA85730.1; PID:9818911
A:Experimental source: strain E2/M3
A>Note: allele 3
A:Genetics: ska
A:Superfamily: streptokinase
C:Keywords: hydrolase; plasminogen activator; virulence

Query Match 100.0%; Score 30; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 6.9; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 1 SVDVEY 6

|||||
Db 17 SVDVEY 22

RESULT 2

AH3460
arsenate reductase [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AH3460
R:DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova.
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Lete
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A>Title: The genome sequence of the facultative intracellular pathogen Brucella melite
A:Reference number: AD3252; PMID:11756688
A:Accession: AH3460
A>Status: preliminary
A:Molecule type: DNA

A;Residues: 1-157 <KUR>
 A;Cross-references: GB:AE008917; PIDN:AA152851.1; PID:g17983693; GSPDB:GN00190
 A;Experimental source: strain 16M
 C;Genetics:
 A;Gene: BME11670
 A;Map position: 1

Query Match 100.0%; Score 30; DB 2; Length 157;
 Best Local Similarity 100.0%; Pred. No. 8.7;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
 |||||
 Db 114 SVDVEY 119

RESULT 3
 T09620
 probable lectin 2 precursor - alfalfa
 C;Species: Medicago sativa (alfalfa)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 C;Accession: T09620
 R;Brill, L.M.; Pietermel, V.R.
 submitted to the EMBL Data Library, March 1998
 A;Description: Legume seed lectin genes: sequence of Molec2 from Alfalfa, Alfalfa and S
 A;Reference number: 216780
 A;Accession: T09620
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-279 <BRI>
 A;Cross-references: EMBL:Y16754
 A;Experimental source: cultivar Chief
 C;Genetics:
 A;Gene: lec2
 C;Superfamily: plant lectin
 C;Keywords: calcium; glycoprotein; lectin
 F;1-25/Domain: signal sequence #status predicted <SIG>
 F;26-279/Product: probable lectin 2 #status predicted <MAT>

Query Match 100.0%; Score 30; DB 2; Length 279;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
 |||||
 Db 208 SVDVEY 213

RESULT 4
 BZSO
 streptokinase (EC 3.4.-.-) - Streptococcus sp.
 C;Species: Streptococcus sp.
 C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 23-Aug-1996
 C;Accession: A00967
 R;Jackson, K.W.; Tang, J.
 Biochemistry 21, 6620-6625, 1982
 A;Title: Complete amino acid sequence of streptokinase and its homology with serine prot
 A;Reference number: A00967; MUID:83127125; PMID:6760891
 A;Accession: A00967
 A;Molecule type: protein
 A;Residues: 1-415 <JAC>
 A;Note: 169-Asp and 181-Asp were also found
 A;Note: this protein is not a protease, but it activates plasminogen by complexing with
 C;Superfamily: streptokinase
 C;Keywords: hydrolase

Query Match 100.0%; Score 30; DB 1; Length 415;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
 |||||
 Db 157 SVDVEY 162

RESULT 5

A22801

streptokinase precursor - Streptococcus "equisimilis"
 C;Species: Streptococcus "equisimilis"

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 24-Sep-1999

C;Accession: A22801

R;Malke, H.; Roe, B.; Ferretti, J.J.

Gene 34, 357-362, 1985

A;Title: Nucleotide sequence of the streptokinase gene from Streptococcus equisimilis 1

A;Reference number: A22801; MUID:85232082; PMID:2989113

A;Accession: A22801

A;Molecule type: DNA

A;Residues: 1-440 <WAL>

A;Cross-references: GB:X72832; NID:g407876; PIDN:CAA51351.1; PID:g407879

A;Experimental source: strain H46A

C;Genetics:

A;Gene: skc

C;Superfamily: streptokinase

Query Match 100.0%; Score 30; DB 1; Length 440;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
 |||||
 Db 183 SVDVEY 188

RESULT 6

S02723

streptokinase G precursor - Streptococcus sp. (strain 19908)

C;Species: Streptococcus sp.

C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 22-Jun-1999

C;Accession: S02723

R;Walter, F.; Siegel, M.; Malke, H.

Nucleic Acids Res. 17, 1262, 1989

A;Title: Nucleotide sequence of the streptokinase gene from a group-G Streptococcus.

A;Reference number: S02723; MUID:89160285; PMID:2922269

A;Accession: S02723

A;Molecule type: DNA

A;Residues: 1-440 <WAL>

A;Cross-references: EMBL:X13400; NID:g47095; PIDN:CAA31766.1; PID:g47096

C;Genetics:

A;Gene: skg

C;Superfamily: streptokinase

F;1-26/Domain: signal sequence #status predicted <SIG>

F;27-440/Product: streptokinase #status predicted <MAT>

Query Match 100.0%; Score 30; DB 2; Length 440;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
 |||||
 Db 183 SVDVEY 188

RESULT 7

S02724

streptokinase A precursor - Streptococcus pyogenes (strain SF130/13)

C;Species: Streptococcus pyogenes

C;Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 22-Jun-1999

C;Accession: S02724

R;Walter, F.; Siegel, M.; Malke, H.

Nucleic Acids Res. 17, 1261, 1989

A;Title: Nucleotide sequence of the streptokinase gene from a Streptococcus pyogenes t

A;Reference number: S02724; MUID:89160284; PMID:2846590

A;Accession: S02724

A;Molecule type: DNA

A;Residues: 1-440 <WAL>

A;Cross-references: EMBL:X13399; NID:g47435; PIDN:CAA31765.1; PID:g47436

C:Genetics:

A:Gene: ska

C:Superfamily: streptokinase

F:1-26/Domain: signal sequence #status predicted <SIG>

P:27-440/Product: streptokinase #status predicted <MAT>

Query Match 100.0%; Score 30; DB 2; Length 440;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6

Db 183 SVDVEY 188

RESULT 8

T10729

transferrin-like protein Ttf-1, salt-induced - green alga (*Dunaliella salina*)C:Species: *Dunaliella salina*

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: T10729

R:Fishier, M.; Gokhman, I.; Pick, U.; Zamir, A.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z17101

A:Accession: T10729

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1274 <PIS>

A:Cross-references: EMBL:U77059; NID:g1684791; PID:g1684792

C:Genetics:

A:Gene: ttf1

C:Superfamily: transferrin repeat homology

Query Match 100.0%; Score 30; DB 2; Length 1274;

Best Local Similarity 100.0%; Pred. No. 89;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6

Db 1186 SVDVEY 1191

RESULT 9

T31951

hypothetical protein B0047.4 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C:Accession: T31951

R:Pauley, A.; Gosia, D.; Ozersky, P.

submitted to the EMBL Data Library, July 1997

A:Description: The sequence of *C. elegans* cosmid B0047.

A:Reference number: Z21102

A:Accession: T31951

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-186 <PAU>

A:Cross-references: EMBL:AF016658; PIDN:AAB66044.1; GSPDB:GN00020; CESP:B0047.4

A:Experimental source: strain Bristol N2, clone B0047

C:Genetics:

A:Gene: CESP:B0047.4

A:Map position: 2

A:Introns: 51/2

Query Match 96.7%; Score 29; DB 2; Length 186;

Best Local Similarity 83.3%; Pred. No. 18;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6

Db 79 SVDVEY 84

RESULT 10

C97297

probable phosphatase, HAD superfamily [imported] - *Clostridium acetobutylicum*C:Species: *Clostridium acetobutylicum*

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: C97297

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee

J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium C

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: C97297

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-215 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK81166.1; PID:g15026304; GSPDB:GN00168

A:Experimental source: *Clostridium acetobutylicum* ATCC824

C:Genetics:

A:Gene: CAC3231

Query Match 96.7%; Score 29; DB 2; Length 215;

Best Local Similarity 83.3%; Pred. No. 22;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6

Db 25 SVDVEY 30

RESULT 11

T16011

hypothetical protein F09F7.3 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999

C:Accession: T16011

R:Pauley, A.

submitted to the EMBL Data Library, May 1994

A:Description: The sequence of *C. elegans* cosmid F09F7.

A:Reference number: Z18445

A:Accession: T16011

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1207 <PAU>

A:Cross-references: EMBL:U00050; NID:g485108; PID:g485110; PIDN:AAA50695.1; CESP:F09F7.

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:F09F7.3

A:Introns: 23/1; 297/3; 600/2; 630/2; 724/3; 789/3; 916/1; 1102/2; 1150/3

C:Superfamily: DNA-directed RNA polymerase 132K polypeptide

Query Match 96.7%; Score 29; DB 2; Length 1207;

Best Local Similarity 83.3%; Pred. No. 1.5e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVDVEY 6

Db 170 SVDIEY 175

RESULT 12

T31082

endo-1,4-beta-xylanase (EC 3.2.1.8) - *Caldicellulosiruptor* sp.C:Species: *Caldicellulosiruptor* sp.

C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000

C:Accession: T31082

R:Morris, D.D.; Gibbs, M.D.; Ford, M.; Thomas, J.; Bergquist, P.L.

submitted to the EMBL Data Library, December 1997

A:Description: Family 10 and 11 xylanase genes from *Caldicellulosiruptor* sp. Rt69B.1.

A:Reference number: Z20972

A:Accession: T31082

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1595 <MOR>

A:Cross-references: EMBL:AF036923; NID:g2760904; PID:g2760908; PIDN:AAB95325.1

C;Genetics:

A;Gene: xynB

C;Keywords: glycosidase; hydrolase

Query Match 96.7%; Score 29; DB 2; Length 1595;

Best Local Similarity 83.3%; Pred. No. 2e+02; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0;

Qy 1 SVDVEY 6

|:|:|

Db 1548 SIDVEY 1553

RESULT 13

IOBS44

N;Alternate names: dnaI protein homolog, 44K; hypothetical protein Y (dnaB 3' region)

C;Species: Bacillus subtilis

C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 19-Jan-2001

C;Accession: B24720; C26580; F69617

R;Ogasawara, N.; Moriya, S.; Mazza, P.G.; Yoshikawa, H.

Nucleic Acids Res. 14, 9989-9999, 1986

A;Title: Nucleotide sequence and organization of dnaB gene and neighbouring genes on the

A;Reference number: A93650; MUID:87117549; PMID:3027671

A;Accession: B24720

A;Molecule type: DNA

A;Residues: 1-311 <OGA>

A;Cross-references: GB:X04963; NID:g39880; PIDN:CAA28633.1; PID:g39881

R;Hoshino, T.; McKenzie, T.; Schmidt, S.; Tanaka, T.; Sueoka, N.

Proc. Natl. Acad. Sci. U.S.A. 84, 653-657, 1987

A;Title: Nucleotide sequence of Bacillus subtilis dnaB: a gene essential for DNA replica

A;Reference number: A94709; MUID:87118226; PMID:3027697

A;Accession: C26580

A;Molecule type: DNA

A;Residues: 1-18, 'N', 20-23, 'T', 25-206 <HOS>

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertea

C.; Bron, S.; Brouillet, S.; Bruchli, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall

tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,

T.; Winter, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, M

A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: F69617

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-311 <KUN>

A;Cross-references: GB:Z99118; GB:AL009126; NID:g2635200; PIDN:CAB14858.1; PID:g2635363

A;Experimental source: strain 168

C;Genetics:

A;Gene: dnaI

C;Superfamily: 44K dnaA protein homolog

C;Keywords: ATP; nucleotide binding; P-loop

F;168-175/Region: nucleotide-binding motif A (P-loop)

Query Match

Best Local Similarity 93.3%; Score 28; DB 1; Length 311;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6

|:|:|

Db 93 SIDIEY 98

RESULT 14

T27619

hypothetical protein ZC504.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C;Accession: T27619

R;Kershaw, J.

submitted to the EMBL Data Library, July 1995

A;Reference number: Z20394

A;Accession: T27619

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-629 <WIL>

A;Cross-references: EMBL:Z50029; PIDN:CAA90341.1; GSPDB:GN00028; CESP:ZC504.2

A;Experimental source: clone ZC504

C;Genetics:

A;Gene: CESP:ZC504.2

A;Map position: X

A;Introns: 11/3; 63/1; 114/3; 166/2; 187/3; 233/3; 348/2; 405/1; 431/2; 473/1; 581/3

C;Superfamily: acetylcholine receptor

Query Match 93.3%; Score 28; DB 2; Length 629;

Best Local Similarity 66.7%; Pred. No. 1.2e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6

|:|:|

Db 200 SIDIEY 205

RESULT 15

A60745

major outer membrane protein - Fusobacterium nucleatum (strain FeV1) (fragment)

C;Species: Fusobacterium nucleatum

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C;Accession: A60745

R;Bakken, V.; Aaro, S.; Jensen, H.B.

J. Gen. Microbiol. 135, 3253-3262, 1989

A;Title: Purification and partial characterization of a major outer-membrane protein o

A;Reference number: A60745; MUID:90257576; PMID:2636259

A;Accession: A60745

A;Molecule type: protein

A;Residues: 1-50 <BAK>

A;Note: sequences of the homologous protein from strains F6, F3, and ATCC 10953 were d

C;Keywords: membrane protein

Query Match

Best Local Similarity 90.0%; Score 27; DB 2; Length 50;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6

|:|:|

Db 38 SVDVQY 43

Search completed: January 10, 2003, 11:47:49

Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 11:45:21 ; Search time 29 Seconds
(without alignments)
42.630 Million cell updates/sec

Title: US-09-919-703-1
Perfect score: 30
Sequence: 1 SVDVEY 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriaph.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	157	16	Q8YF57
2	30	100.0	279	10	Q49899
3	30	100.0	414	2	Q53284
4	30	100.0	891	5	Q25177
5	30	100.0	1274	10	P93125
6	29	96.7	60	5	Q9NFR4
7	29	96.7	74	16	Q9L5P9
8	29	96.7	186	5	O16567
9	29	96.7	215	16	Q97E84
10	29	96.7	520	5	Q9WSV6
11	29	96.7	547	5	Q8WSF9
12	29	96.7	615	3	Q9P861
13	29	96.7	742	5	Q8SZV9
14	29	96.7	830	2	Q9AGP8
15	29	96.7	836	10	Q9FGN0
16	29	96.7	1154	5	Q27492
					Q8YF57 brucella me
					Q49899 medicago sa
					Q53284 streptococc
					Q25177 hydra atten
					P93125 dunaliella
					Q9NFR4 drosophila
					Q9L5P9 salmonella
					O16567 caenorhabdi
					Q97E84 clostridium
					Q9WSV6 drosophila
					Q8WSF9 aplysia cal
					Q9P861 arxula aden
					Q8SZV9 drosophila
					Q9AGP8 arthrobacte
					Q9FGN0 arabidopsis
					Q27492 caenorhabdi

17	29	96.7	1190	5	Q9VAD6
18	29	96.7	1595	2	O52373
19	28	93.3	146	9	Q8SCX2
20	28	93.3	148	2	Q9ADW6
21	28	93.3	433	12	Q8V3P9
22	28	93.3	579	5	Q23355
23	28	93.3	907	5	Q9U3Q5
24	27	90.0	136	12	O12377
25	27	90.0	180	5	Q09968
26	27	90.0	271	16	Q8X2B1
27	27	90.0	271	16	Q8X298
28	27	90.0	275	10	Q9FWP3
29	27	90.0	277	17	Q9HKV3
30	27	90.0	302	12	Q933F5
31	27	90.0	307	17	Q8ZTC7
32	27	90.0	349	16	Q9L074
33	27	90.0	355	2	Q48587
34	27	90.0	355	2	P71435
35	27	90.0	368	2	Q47904
36	27	90.0	368	2	Q47903
37	27	90.0	368	16	Q8RHY1
38	27	90.0	370	2	Q47905
39	27	90.0	372	2	Q47913
40	27	90.0	375	16	Q8X4C8
41	27	90.0	391	3	Q12132
42	27	90.0	391	12	O89268
43	27	90.0	399	16	O05452
44	27	90.0	405	17	Q8U2Q0
45	27	90.0	407	16	Q8X5F6

ALIGNMENTS

RESULT 1

Q8YF57 Q8YF57 PRELIMINARY; PRT; 157 AA.
AC Q8YF57;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Arsenate reductase.
GN EMBL1670.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / BIOTYPE 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mujar C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA Seikov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AB009602; AAL52851.1; -;
DR InterPro; IPR000106; Low_mwt_PTPase.
DR Pfam; PF01451; LMWPC; 1.
DR SMART; SM00226; LMWPC; 1.
KW Complete proteome.
SQ SEQUENCE 157 AA; 17446 MW; D8AA7748107ECB65 CRC64;

Query Match 100.0%; Score 30; DB 16; Length 157;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6

Db 114 SVDVEY 119

RESULT 2
O49899 PRELIMINARY; PRT; 279 AA.
AC O49899;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Lactin precursor.
GN LEC2.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
OX NCBI_TaxID=3879;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=CV, CHIEF;
RC Brill L.M., Pieterse V.R.;
RA "Legume seed lectin genes: sequence of Mslec2 from Alfalfa, Alfalfa
RT and Sweet Clover southern blot analysis, and stable transformation of
RT Alfalfa with antisense-lectin constructs.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RR EMBL; Y16754; CAA76366.1; -
DR HSP; P04122; ILOE.
DR InterPro; IPR000985; Lactin_legA.
DR InterPro; IPR001220; Lactin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lactin_legA; 1.
DR ProDom; PD000711; Lactin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Signal.
FT SIGNAL.
FT CHAIN.
FT SIGNAL.
FT CHAIN.
SQ SEQUENCE 279 AA; 30551 MW; 1315F022BABDA360 CRC64;
Query Match 100.0%; Score 30; DB 10; Length 279;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVDVEY 6
Db 208 SVDVEY 213

RESULT 3
O53284 PRELIMINARY; PRT; 414 AA.
AC Q53284;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SKC-2.
OS Streptococcus equisimilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=119602;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=93000998; PubMed=1368792;
RX Estrada M.P., Hernandez L., Perez A., Rodriguez P., Serrano R.,
RA Rubiera R., Pedraza A., Padron G., Antuch W., de la Fuente J.,
RA Herrera L.;
RT "High level expression of streptokinase in Escherichia coli.";
RL Biotechnology 10:1138-1142(1992).
DR EMBL; S46536; AAC60418.1; -
DR InterPro; IPR004093; Staphylokinase.
DR Pfam; PF02821; Staphylokinase; 3.
SQ SEQUENCE 414 AA; 47254 MW; F75B85831B766904 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVDVEY 6
Db 157 SVDVEY 162

RESULT 4
Q25177 PRELIMINARY; PRT; 891 AA.
AC Q25177;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Receptor protein-tyrosine kinase (Fragment).
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=UCI;
RC Chen R.L., Steele R.E.;
RA Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U24116; AAG5223.1; -
DR HSP; P11362; IFGK.
DR InterPro; IPR002106; AATRNA_ligaseII.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; ig; 2.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00103; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00410; IG_like; 2.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Kinase; Tyrosine-protein Kinase.
FT NON TER 891
SQ SEQUENCE 891 AA; 100976 MW; 1CFDP48CCA298176 CRC64;
Query Match 100.0%; Score 30; DB 5; Length 891;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SVDVEY 6
Db 137 SVDVEY 142

RESULT 5
P93125 PRELIMINARY; PRT; 1274 AA.
AC P93125;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Ttf-1.
GN TTF1.
OS Drosophila salina.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Dunaliellaceae; Dunaliella.
OX NCBI_TaxID=3046;
[1]
RN SEQUENCE FROM N.A.
RP Fisher M., Gokhman I., Pick U., Zamir A.;
RA Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; U77059; AAB36531.1; -

```

DR HSP; P56410; 1A0V.
DR InterPro; IPR001156; Transferrin.
DR Pfam; PF00405; transferrin; 5.
DR SMART; SMO0094; TR FER; 1.
SQ SEQUENCE 1274 AA; 136668 MW; 6C2EE9D914097699 CRC64;

Query Match 100.0%; Score 30; DB 10; Length 1274;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
Db 1186 SVDVEY 1191

RESULT 6
Q9NFR4 PRELIMINARY; PRT; 60 AA.
AC Q9NFR4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha 4 subunit (Fragment).
GN NACR-ALPHA-80B OR NACRALPHA-80B OR CG12414 OR CG17552.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Lansdell S.J., Millar N.S.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
DR EMBL; AJ272160; CAB77446.1; -.
DR FlyBase; FBgn0037212; nacr-alpha-80B.
DR InterPro; IPR001175; Neur channel.
DR Pfam; PF02931; Neur chan IBD; 1.
DR PROSITE; PS00236; NEUROPE ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
FT NON TER 1
FT TER 60
SQ SEQUENCE 60 AA; 6868 MW; E3A9483419FAD0C6 CRC64;

Query Match 96.7%; Score 29; DB 5; Length 60;
Best Local Similarity 83.3%; Pred. No. 17;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
Db 34 SVDVEY 39

RESULT 7
Q9LSP9 PRELIMINARY; PRT; 74 AA.
AC Q9LSP9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Orf, hypothetical protein (Hypothetical 8.2 kDa protein).
GN R0092 OR HCM1.13C.
OS Salmonella typhi.
OG Plasmid R27, and Plasmid pHCMI.
OC Bacteria; Proteobacteria; Gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC PLASMID=R27;
RX MEDLINE=20280091; PubMed=10773089;

Sherburne C.K., Lawley T.D., Gilmour M.W., Blattner F.R., Burland V.,
Grotbeck E., Rose D.J., Taylor D.E.;
"The complete DNA sequence and analysis of R27, a large IncHI plasmid
from Salmonella typhi that is temperature sensitive for transfer.";
Nucleic Acids Res. 28:2177-2186(2000).
[2]
SEQUENCE FROM N.A.
STRAIN=CT18; PLASMID=PHCM1;
MEDLINE=21534947; PubMed=11677608;
Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jegeris K.,
Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
Whitehead S., Barrall B.G.;
"Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
Nature 413:848-852(2001).
RL Nature 413:848-852(2001).
DR EMBL; AF250878; AAF69930.1; -.
DR EMBL; AL513383; CAD09627.1; -.
KW Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 74 AA; 8199 MW; DCC7A8A73B746EAF CRC64;

Query Match 96.7%; Score 29; DB 16; Length 74;
Best Local Similarity 83.3%; Pred. No. 21;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVDVEY 6
Db 30 SVDVEY 35

RESULT 8
O16567 PRELIMINARY; PRT; 186 AA.
AC O16567;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE B0047.4 protein.
GN B0047.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,
Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fulton L.,
Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
Smaldon M., Smith A., Sonhammer E., Staden R., Sulston J.,
Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
"2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans.";
Nature 368:32-38(1994).
[2]
SEQUENCE FROM N.A.
STRAIN=BRISTOL N2;
RA Pauley A., Goela D., Ozersky P.;
RT "The sequence of C. elegans cosmid B0047.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;

```

Waterston R.,
 Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 EMBL: AF016658; AAB66044.1; -
 InterPro: IPR002083; MATH.
 Pfam: PF00917; MATH; 1.
 SMART: SM00061; MATH; 1.
 SQUENCE 186 AA; 21527 MW; 9E216690AB0E1469 CRC64;
 Query Match 96.7%; Score 29; DB 5; Length 186;
 Best Local Similarity 83.3%; Pred. No. 53;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 1 SVDVEY 6
 79 SIDVEY 84
 -7E84
 -SULT 9
 Q97E84 PRELIMINARY; PRT; 215 AA.
 Q97E84
 01-OCT-2001 (TrEMBLrel. 18, Created)
 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 Predicted phosphatase, HAD superfamily.
 CAC3231.
 Clostridium acetobutylicum.
 Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 Clostridiales; Clostridiaceae; Clostridium.
 NCBI_TaxID=1488;
 [1]
 SQUENCE FROM N.A.
 STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 MEDLINE=21359325; PubMed=11466286;
 Neolling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 Tatunov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 Bennett G.N., Koonin E.V., Smith D.R.;
 "Genome sequence and comparative analysis of the solvent-producing
 bacterium Clostridium acetobutylicum."
 J. Bacteriol. 183:4823-4838(2001).
 EMBL: AE007819; AAK81166.1; -
 InterPro: IPR001454; Hlgase/hydrolase.
 Pfam: PF00702; Hydrolase; 1.
 PRINTS: PR00413; HADHALOGNASE.
 Complete proteome.
 SQUENCE 215 AA; 24421 MW; C4334F6313CBFC15 CRC64;
 Query Match 96.7%; Score 29; DB 16; Length 215;
 Best Local Similarity 83.3%; Pred. No. 62;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 1 SVDVEY 6
 25 SIDVEY 30
 -WSV6
 -SULT 10
 Q9WSV6 PRELIMINARY; PRT; 520 AA.
 Q9WSV6; Q9WSV5;
 01-MAY-2000 (TrEMBLrel. 13, Created)
 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 Putative nicotinic acetylcholine receptor alpha 80B (Fragment).
 NACR-ALPHA-80B OR CG12414 OR CG17552.
 Drosophila melanogaster (Fruit fly).
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 [1]
 SQUENCE FROM N.A.

STRAIN=BERKELEY;
 MEDLINE=20196006; PubMed=10731132;
 Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Hartis N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of Drosophila melanogaster."
 Science 287:2185-2195(2000).
 RL -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNELS FAMILY.
 CC EMBL: AE002665; AAF45409.1; -
 DR FlyBase; FBgn0037212; nacr-alpha-80B.
 DR InterPro; IPR000188; GABA_A_receptor.
 DR InterPro; IPR001175; Neur_channel.
 DR Pfam; PF02932; Neur_chan_LBD; 1.
 DR Pfam; PF02931; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 2.
 KW Glycoprotein; Hypothetical protein; Ionic channel;
 KW Postsynaptic membrane; Transmembrane.
 FT NON_TER 1
 SQ SQUENCE 520 AA; 59339 MW; 5B2AE734E78E7D4D CRC64;
 Query Match 96.7%; Score 29; DB 5; Length 520;
 Best Local Similarity 83.3%; Pred. No. 1.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVDVEY 6
 DB 182 SIDVEY 187
 RESULT 11
 Q9WSF9
 ID Q9WSF9 PRELIMINARY; PRT; 547 AA.
 AC Q9WSF9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Non-alpha nicotinic acetylcholine receptor subunit.
 OS Aplysia californica (California sea hare).

OC Eukaryota; Metazoa; Mollusca; Gastropoda; Opisthobranchia; Anaspidae;
 OC Aplysiidae; Aplysia.
 OX NCBI_TaxID=6500;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sadreyev R.I., Panchin Y.V., Meleshkevich E.A., Moroz L.L.;
 RT "Non-alpha nicotinic acetylcholine receptor subunit from the
 individual identified neurons of Aplysia californica.";
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF322876; AAL37250.1; -;
 DR InterPro; IPR000188; GABAA_receptor.
 DR InterPro; IPR001175; Neur_Channel.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR TIGRFAMs; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUOTR_ION_CHANNEL; UNKNOWN_1.
 KW Receptor.
 SQ SEQUENCE 547 AA; 62535 MW; 3D3F011D78605C91 CRC64;
 Query Match 96.7%; Score 29; DB 5; Length 547;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVDVEY 6
 Db 155 SVDIEY 160
 RESULT 12
 ID Q9P861 PRELIMINARY; PRT; 615 AA.
 AC Q9P861;
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE Ferro-O2-oxidoreductase precursor.
 GN AFET3.
 OS Arxula adeninivorans.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetaceae; Dipodascaceae; mitosporic Dipodascaceae; Arxula.
 OX NCBI_TaxID=37620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=LS3;
 RA Martmann T., Stephan U.W., Bube I., Boeer E., Melzer M.,
 RA Manteuffel R., Stoltenburg R., Kunze G.;
 RT "Post-translational modifications of the AFET3 gene product - a
 component of iron transport system in budding cells and mycelia of the
 yeast Arxula adeninivorans LS3.";
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ277833; CAB90817.1; -;
 DR HSP; P37064; IAOZ.
 DR InterPro; IPR001117; Cu-oxidase.
 DR InterPro; IPR002355; MultiCu_oxidase2.
 DR Pfam; PF00394; Cu-oxidase; 3.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; 1.
 DR PROSITE; PS00080; MULTICOPPER_OXIDASE2; 1.
 KW Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 615 FERRO-O2-OXIDOREDUCTASE.
 SQ SEQUENCE 615 AA; 69624 MW; 7920D42A28715B4 CRC64;
 Query Match 96.7%; Score 29; DB 3; Length 615;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVDVEY 6
 Db 311 SVDIEY 316

RESULT 13

Q8SZV9
 ID Q8SZV9 PRELIMINARY; PRT; 742 AA.
 AC Q8SZV9;
 DT 01-JUN-2002 (T-EMBLrel. 21, Created)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE RE01016p.
 GN CG7880.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY069758; AAL39903.1; -;
 SQ SEQUENCE 742 AA; 84075 MW; 85998EF365194E39 CRC64;
 Query Match 96.7%; Score 29; DB 5; Length 742;
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVDVEY 6
 Db 682 SVDIEY 687
 RESULT 14
 ID Q9AGP8 PRELIMINARY; PRT; 830 AA.
 AC Q9AGP8;
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
 DE N,N-dimethylglycine oxidase.
 GN DMG.
 OS Arthrobacter globiformis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Micrococcales; Micrococcaceae; Arthrobacter.
 OX NCBI_TaxID=1685;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NRRL B-2979;
 RA Meskys R., Harris R.J., Casate V., Baaran J., Scrutton N.S.;
 RT "Genetic organization of the genes involved in dimethylglycine and
 sarcosine degradation in Arthrobacter spp.: implications for glycine
 betaine catabolism.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF329477; AAK16482.1; -;
 DR InterPro; IPR000927; D_aa_oxidase.
 DR InterPro; IPR002536; GCV_T.
 DR InterPro; IPR000205; NAD_binding.
 DR InterPro; IPR000594; Thif_domain.
 DR Pfam; PF01266; DAO; 1.
 DR Pfam; PF01571; GCV_T; 1.
 SQ SEQUENCE 830 AA; 89984 MW; 13DE3C4B3DF325DA CRC64;
 Query Match 96.7%; Score 29; DB 2; Length 830;
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVDVEY 6
 Db 801 SVDIEY 806

```

RESULT 15
Q9FGN0 PRELIMINARY; PRT; 836 AA.
AC Q9FGN0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similarity to unknown protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB025621; BAB09754.1; -. A875BAE2F18F61BD CRC64;
SQ SEQUENCE 836 AA; 92259 MW; A875BAE2F18F61BD CRC64;

Query Match 96.7%; Score 29; DB 10; Length 836;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SDVEY 6
Db 772 SVDIEY 777

```

Search completed: January 10, 2003, 11:47:27
 Job time : 31 secs